GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 30, 2003, 22:32:51 ; Search time 2816 Seconds
(without alignments)
2966.089 Million cell updates/sec

Title: US-09-925-637-64
Perfect score: 287
Sequence: 1 MTNLPMNKLIDEVNNELSVA.....ELTQIDEQFNTKHLLEIVDL 287

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

arched: 2054640 seqs, 14551402878 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4106490

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=x1h
-Q=/cgn2_1/USPTO_spool/US09925637/runat_23052003_174850_24112/app_query.fasta_1.455
-DB=GenEmbl -QFMT=fastap -SUFFIX=oli.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09925637_@CGN_1_1_2496_@runat_23052003_174850_24112 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	197	68.6	333750	1	AP004827	AP004827 Staphyloc
c	2	132	46.0	301550	1	AP003134	AP003134 Staphyloc
c	3	132	46.0	346900	1	AP003362	AP003362 Staphyloc
	4	102	35.5	882	6	AR106466	AR106466 Sequence
	5	37	12.9	783	6	AX141501	AX141501 Sequence
	6	37	12.9	3444	1	AF270104	AF270104 Staphyloc
	7	37	12.9	3444	6	AX145422	AX145422 Sequence
c	8	37	12.9	3454	1	AF269958	AF269958 Staphyloc
c	9	37	12.9	3454	6	AX145276	AX145276 Sequence
c	10	37	12.9	4045	1	AF269889	AF269889 Staphyloc
c	11	37	12.9	4045	6	AX145207	AX145207 Sequence
	12	33	11.5	486	6	AR097529	AR097529 Sequence
	13	33	11.5	486	6	AR142282	AR142282 Sequence
	14	33	11.5	486	6	E15319	E15319 Micrococcus
	15	33	11.5	1080	1	AB003187	AB003187 Micrococc
	16	27	9.4	2268	1	LMO012349	AJ012349 Listeria
	17	27	9.4	5635	6	AX416832	AX416832 Sequence
	18	27	9.4	250050	1	AL591978	AL591978 Listeria
	19	19	6.6	957	1	AB016093	AB016093 Synechoco
c	20	19	6.6	195269	6	AX417035	AX417035 Sequence
c	21	19	6.6	300600	1	AP005369	AP005369 Thermosyn
	22	19	6.6	333050	1	AL596168	AL596168 Listeria
	23	19	6.6	349980	6	AX417044	AX417044 Sequence
	24	18	6.3	10029	1	AE006320	AE006320 Lactococc
	25	18	6.3	10029	1	AE013090	AE013090 Thermoana
c	26	18	6.3	10165	1	AE006583	AE006583 Streptoco
c	27	18	6.3	13536	1	AE010066	AE010066 Streptoco
c	28	18	6.3	50463	1	AE014159	AE014159 Streptoco
	29	18	6.3	348050	1	AP003581	AP003581 Nostoc sp
	30	16	5.6	155932	1	CJ11168X6	AL139079 Campyloba
	31	15	5.2	891	6	AX394976	AX394976 Sequence
	32	15	5.2	891	6	AX398494	AX398494 Sequence
	33	15	5.2	891	6	AX398804	AX398804 Sequence
	34	15	5.2	4451	1	WSAJ3049	AJ003049 Wolinella
	35	14	4.9	811	6	AX432600	AX432600 Sequence
	36	14	4.9	894	6	AR148386	AR148386 Sequence
	37	14	4.9	894	6	E05337	E05337 DNA encodin
	38	14	4.9	894	6	E11866	E11866 Mutated DNA
	39	14	4.9	894	6	E11867	E11867 Mutated DNA
	40	14	4.9	894	6	E11869	E11869 Mutated DNA
	41	14	4.9	894	6	E14762	E14762 gDNA encodi
	42	14	4.9	894	6	E16095	E16095 DNA encodin
	43	14	4.9	894	6	E27521	E27521 Geranyl dip
	44	14	4.9	894	6	E54864	E54864 Process for
	45	14	4.9	1260	1	BACFDPS	D13293 B. stearoth

ALIGNMENTS

RESULT 1

```
AP004827/c
LOCUS       AP004827              333750 bp    DNA     linear   BCT 02-JUL-2002
DEFINITION  Staphylococcus aureus subsp. aureus MW2 DNA, complete genome,
            strain:MW2, section 6/10.
ACCESSION   AP004827 BA000033
VERSION     AP004827.1  GI:21204509
KEYWORDS    .
SOURCE      Staphylococcus aureus subsp. aureus MW2 (strain:MW2) DNA.
  ORGANISM  Staphylococcus aureus subsp. aureus MW2
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Baba,T., Takeuchi,F., Kuroda,M., Yuzawa,H., Aoki,K., Oguchi,A.,
            Nagai,Y., Iwama,N., Asano,K., Naimi,T., Kuroda,H., Cui,L.,
            Yamamoto,K. and Hiramatsu,K.
  TITLE     Genome and virulence determinants of high virulence
            community-acquired MRSA
  JOURNAL   Lancet 359 (9320), 1819-1827 (2002)
  MEDLINE   22040717
   PUBMED   12044378
   ERENCE   2  (bases 1 to 333750)
   UTHORS   Aoki,K., Oguchi,A., Nagai,Y., Asano,K., Iwama,N., Baba,T.,
            Kuroda,M., Hiramatsu,K. and Kikuchi,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAR-2002) Akio Oguchi, National Institute of
            Technology and Evaluation, Biotechnology Center; 2Chome 49-10
            Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
            (E-mail:oguchi@nite.go.jp, URL:http://www.bio.nite.go.jp/,
            Tel:81-3-3481-8423, Fax:81-3-3481-8424)
FEATURES             Location/Qualifiers
     source          1. .333750
                     /organism="Staphylococcus aureus subsp. aureus MW2"
                     /strain="MW2"
                     /db_xref="taxon:196620"
     gene            complement(278. .619)
                     /gene="MW1341"
     CDS             complement(278. .619)
                     /gene="MW1341"
                     /note="ORFID:MW1341"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAB95206.1"
                     /db_xref="GI:21204510"
                     /translation="MTNITKEVFDNLEQEIDVFAKNKTLGSSEAKPYLDEYHSKIIDY
                     FKQVNDITGNIDFDKLNQYPVVPMNFKERYDYMIERKYHFMGYRQMKTFKTELIKMNA
                     SYQTRLKNKQV"
     gene            complement(624. .1283)
                     /gene="nth"
     CDS             complement(624. .1283)
                     /gene="nth"
                     /note="ORFID:MW1342"
                     /codon_start=1
                     /transl_table=11
                     /product="endonuclease-like protein"
                     /protein_id="BAB95207.1"
                     /db_xref="GI:21204511"
                     /translation="MVSKKKALEMIDVIANMFPDAECELKHDNPFELTIAVLLSAQCT
                     DVLVNRVTTELFKKYKTPEDYLAVSDEELMNDIRSIGLYRNKAKNIKKLCQSLIDQFN
                     GEIPQTHKELESLAGVGRKTANVVMSVAFDEPSLAVDTHVERVSKRLGINRWKDNVRQ
                     VEDRLCSVIPRDRWNRSHHQLIFFGRYHCLARKPKCDICPLLEDCREGQKRYKASLKE
                     A"
     gene            complement(1273. .1959)
                     /gene="MW1343"
     CDS             complement(1273. .1959)
                     /gene="MW1343"
                     /note="ORFID:MW1343
                     hypothetical protein, similar to chromosome replication
                     initiation protein dnaD"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB95208.1"
                     /db_xref="GI:21204512"
                     /translation="MDKYQLKARPVVIRRELLDHYSDLGLDEQDLVILLKLIYASETS
                     NKQPSIELLQKGSTMQPRDITMVIQNLIQRELLELQVQKDEEGRFTEYMNLDPFFEKL
                     SHILKQQSMETKEQNSKEKFKQLFRVLEDTFARPLSPYEIETLNQWIDVDKHDSAIIQ
                     AALDEANSLNKLSFKYMDRILLNWKKNNVKTIDDSRKIREKFNKPKMTHTVKTVPKFD
                     WLNGENLDGK"
     gene            complement(2287. .3579)
                     /gene="asnS"
     CDS             complement(2287. .3579)
                     /gene="asnS"
                     /note="ORFID:MW1344"
                     /codon_start=1
                     /transl_table=11
                     /product="asparaginyl-tRNA synthetase"
                     /protein_id="BAB95209.1"
                     /db_xref="GI:21204513"
                     /translation="MKTTIKQAKDHLNQDVTIGAWLTNKRSSGKIAFLQLRDGTGFMQ
                     GVVVKSEVDEEVFKLAKEITQESSLYVTGTITEDNRSDLGYEMQVKSIEVISEAHDYP
                     ITPKNHGTEFLMDHRHLWLRSKKQHAVMKIRNEVIRATYEFFNKDGFTKVDPPILTAS
                     APEGTSELFHTKYFDQDAFLSQSGQLYLEAAAMAHGKVFSFGPTFRAEKSKTRRHLIE
                     FWMIEGEMAFTNHAESLEIQEQYVTHVVKSVLENCKLELKILERDTSKLEKVATPFPR
                     ISYDDAIEFLKSEGFDDIEWGEDFGAPHETAIANHYDLPVFITNYPTKIKPFYMQPNP
                     ENEETVLCADLIAPEGYGEIIGGSERVDDLELLEQRVKEHGLDEEAYSYYLDLRRYGS
                     VPHCGFGLGLERTVAWISGVEHVRETAPFPRLLNRLYP"
     gene            complement(3901. .6594)
                     /gene="dinG"
     CDS             complement(3901. .6594)
                     /gene="dinG"
                     /note="ORFID:MW1345"
                     /codon_start=1
                     /transl_table=11
                     /product="probable ATP-dependent DNA helicase dinG"
                     /protein_id="BAB95210.1"
                     /db_xref="GI:21204514"
                     /translation="MGMATYAVVDLETTGNQLDFDDIIQIGITFVRNNQIIDTYHSMI
                     RTNLEIPPFIQALTSIEENMLQQAPYFNQVAQEIYDKIKDCIFVAHNVDFDLNFIKKA
                     FKDCNIQYRPKKVIDTLEIFKIAFPTDKSYQLSELAEAHGITLANAHRADEDAATTAK
                     LMILAFEKFEKLPLDTLKQLYYLSKQLKYDLYDIFFEMVRQYDAKPLDKSYEKFEQII
                     YRKQVDFKKPTTNYNGSLKSLYSKAVDQLGLTYRPQQLYLAETILDQLMHSEKAMIEA
                     SLGSGKSLAYLLAALMYNIETGKHVMISTNTKLLQSQLLEKDIPAMNEALNFKINALL
                     IKSKSDYISLGLISQILKDDTSNYEVNILKMQLLIWITETPSGDIQELNLKGGQKMYF
                     DQKIETYVPARHDVHYYNFIKRNAQNIQIGITNHAHLIHSDVENSIYQLFDDCIVDEA
                     HRLPDYALNQVTNELSYADIKYQLGLIGKNENEKLLKAIDQLEKQRILEKLDIAPIDI
                     FGLKASMNEIHELNEQLFSTIFTIINDSDVYDDDIHRFHNVFTFETKDILKDLHAIID
                     KLNKTLEIFNGISHKTVKSLRKQLLYLKDKFKNIEQSLKAGHTSFISIKNLSQKSTIR
                     LYVKDYAVKDVLTKQVLEKFKSLIFISGTLKFNHSFEAFKQLFNKDVHFNTFEVNTSL
                     QSAKNTSVFIPSDVASYQYKNIDEYVASIVSYIIEYTTITSSKCLVLFTSYKMMHMVQ
                     DMLNELPEFEDYVVLTQQQNQNYKIVQQFNNFDKAILLGTSTFFEGFDFQANGIKCVM
                     IAKLPFMNKHNAKYWLMDSEFTSTFKEYVLPDAVTRFRQGLGRLIRSENDRGIIVSFD
                     DRLINSNYKNFFEQTLENYRQKKGDIQQFGKLLRQIQKKKK"
     gene            complement(6618. .7589)
                     /gene="MW1346"
     CDS             complement(6618. .7589)
                     /gene="MW1346"
                     /note="ORFID:MW1346"
                     /codon_start=1
                     /transl_table=11
                     /product="truncated hypothetical protein"
                     /protein_id="BAB95211.1"
                     /db_xref="GI:21204515"
                     /translation="MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLK
                     LEGCKIDSVNHKGHLLQQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSL
                     VGNQSSFFILSDEQTKGRGRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALG
                     IRDAIQHFSQDEVKVKWPNDIYIDNGKVCGFLTEMVANNDGIEAIICGIGINLTQQLE
                     NFDESIRHRATSVQLHDKNKLDRYQFLERLLQEIEKRYNQFLTLPFSEIREEYIAASN
                     IWNRTLLFTENDKQFKGQAIDLDYDGYLIVRDEAGESHRLISADIDF"
     gene            complement(7576. .8778)
                     /gene="MW1347"
     CDS             complement(7576. .8778)
                     /gene="MW1347"
                     /note="ORFID:MW1347
                     hypothetical protein, similar to poly(A) polymerase"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB95212.1"
```

```
                     /db_xref="GI:21204516"
                     /translation="MDKSLFEQARPILEQIQDNGFEAYYVGGSVRDYVMGRNIHDIDI
                     TTSATPDEIESIFSHTIPVGKEHGTINVVFNDENYEVTTFRAEEDYVDHRRPSGVTFV
                     RDLYEDLQRRDFTMNAIAMDTAYKLYDYFDGQQDINNRIIRTVGIAEERFQEDALRMI
                     RCLRFQSQLSFDIATETFEAMRIQMADIKFLSIERIVIELTKLMRGINVEKSFNHLKS
                     LKAFNYMPYFEHLDMNQINVTEAIDLELLIAIVSVKFDINYSLKPLKLSNRQVKDINQ
                     YIQIMNALPSIITKEQLKMFVYDYDTHLIKNVMVAADVIKANDIQGHEPLIVNLQTID
                     ETLHRLPMHNRKDMMVNGGVLMAHLNAKSGPWLKDVLRQIEIAIVTGKVSNEETEILK
                     WVDNHVKI"
     gene            complement(8783. .9925)
                     /gene="MW1348"
     CDS             complement(8783. .9925)
                     /gene="MW1348"
                     /note="ORFID:MW1348
                     hypothetical protein, similar to lipopolysaccharide
                     biosynthesis-related pr homolog"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB95213.1"
                     /db_xref="GI:21204517"
                     /translation="MKIGITCYPSMGGSGIIATELGIKLAERGHEVHFITSNIPFRIR
                     KPLPNMIFHQVEVNQYAVFQYPPYDITLSTKIAEVIKEYGLDLLHMHYAVPHAICGIL
                     AREMSGKDIKIMTTLHGTDITVLGYDHSLQGAIKFGIEKSDIVTSVSKSLAQETHEII
                     ETNKEIIPIYNFVRENEFPTKHNTALKSQFGIAPDEKVLIHVSNFRQVKRIDTIIETF
                     AKVREKIPSKLILLGDGPELVPMRQLTKELNVEEDVLFLGKQDCVSEFYQLSDLVLLL
                     SEKESFGLTLLEAMKTGVVPIGSNAGGIKEVIKHGETGFVVDVGDCDSASDYAIRLLE
                     DKVLYNKLQKNMLADIAERFGSELITDQYEYYYQKMLNEHNKSKGE"
     gene            complement(10167. .10484)
                     /gene="MW1349"
     CDS             complement(10167. .10484)
                     /gene="MW1349"
                     /note="ORFID:MW1349"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB95214.1"
                     /db_xref="GI:21204518"
                     /translation="MKSMVEMQREVDEYIGQFKTGYFSPLANLARLTEEVGELAREIN
                     HTYGEKKKDSEEANTIKAELGDNLFVLLCLANSMGIDMTESFNETMEKFNTRDKNRF
                     ERK"
     gene            complement(10820. .11518)
                     /gene="MW1350"
     CDS             complement(10820. .11518)
                     /gene="MW1350"
                     /note="ORFID:MW1350"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB95215.1"
                     /db_xref="GI:21204519"
                     /translation="MSSIFTMIIYFVILMVLPLWAQHKVKSNYEKYSQVRSTSGKTGR
                     EVALEILHANGIYDVEVVKGEGFLTDHYDPKKKVVSLSPANYDRPSVAGTAIAAHEVG
                     HAIQDHQGYWFLRFRAALVPVANLGSSLSYMIIMLGIILTAIGSAFGSTALWIGAGLM
                     SLAVLFSIVTLPVEFDASSRAMKQITALNIVNEKEYKHARKVLSAAAMTYVAATAVAV
                     AELVRIILLARSSD"

Alignment Scores:
Pred. No.:              5.48e-195     Length:          333750
Score:                  197.00        Matches:         286
Percent Similarity:     99.31%        Conservative:    0
Best Local Similarity:  99.31%        Mismatches:      1
Query Match:            68.64%        Indels:          2
DB:                     1             Gaps:            0

US-09-925-637-64 (1-287) x AP004827 (1-333750)

Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118374 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 118315

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118314 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 118255

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118254 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 118195

Qy         61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118194 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 118135

Qy         81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118134 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 118075

Qy        101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118074 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 118015

Qy        121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118014 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 117955

Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117954 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 117895

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117894 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 117835

Qy        181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThr-ThrLy 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117834 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGC-AACTAA 117776

Qy        200 sGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLe 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117775 AGAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTT 117716

Qy        220 uLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAs 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117715 ATTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAA 117656

Qy        240 nAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTy 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117655 TAATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTA 117596

Qy        260 rHisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHi 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117595 TCATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACA 117536

Qy        280 sLeuLeuGluIleValAspLeu 287
              |||||||||||||||||||||||
Db     117535 CTTATTAGAAATCGTTGATTTA 117514

RESULT 2
AP003134/c
LOCUS       AP003134               301550 bp    DNA     linear   BCT 02-JUL-2002
DEFINITION  Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
            genome, section 6/10.
ACCESSION   AP003134 BA000018
VERSION     AP003134.2  GI:14349226
KEYWORDS    .
SOURCE      Staphylococcus aureus subsp. aureus N315 (sub_species:aureus N315,
            strain:N315) DNA.
  ORGANISM  Staphylococcus aureus subsp. aureus N315
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
            Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M.,
            Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
            Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
            Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
            Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
            Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
```

```
  TITLE     Whole genome sequencing of meticillin-resistant Staphylococcus
            aureus
  JOURNAL   Lancet 357 (9264), 1225-1240 (2001)
  MEDLINE   21311952
   PUBMED   11418146
REFERENCE   2  (bases 1 to 301550)
  AUTHORS   Aoki,K., Oguchi,A., Hosoyama,A., Nagai,Y., Kuroda,M., Hiramatsu,K.
            and Kikuchi,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JAN-2001) Akio Oguchi, National Institute of
            Technology and Evaluation, Biotechnology Center; 2Chome 49-10
            Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
            (E-mail:oguchi@nite.go.jp, URL:http://www.bio.nite.go.jp/,
            Tel:81-3-3481-8423, Fax:81-3-3481-8424)
COMMENT     On Jun 12, 2001 this sequence version replaced gi:13701258.
FEATURES             Location/Qualifiers
     source          1. .301550
                     /organism="Staphylococcus aureus subsp. aureus N315"
                     /strain="N315"
                     /sub_species="aureus N315"
                     /db_xref="taxon:158879"
     gene            complement(249. .947)
                     /gene="SA1293"
     CDS             complement(249. .947)
                     /gene="SA1293"
                     /note="ORFID:SA1293"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB42553.1"
                     /db_xref="GI:13701259"
                     /translation="MSSIFTMIIYFVILMVLPLWAQHKVKSNYEKYSQVRSTSGKTGR
                     EVALEILHANGIYDVEVVKGEGFLTDHYDPKKKVVSLSPANYDRPSVAGTAIAAHEVG
                     HAIQDHQGYWFLRFRAALVPVANLGSSLSYMIIMLGIILTAIGSAFGSTALWIGAGLM
                     SLAVLFSIVTLPVEFDASSRAMKQITALNIVNEKEYKHARKVLSAAAMTYVAATAVAV
                     AELVRIILIARSSD"
     gene            complement(1001. .1588)
                     /gene="SA1294"
     CDS             complement(1001. .1588)
                     /gene="SA1294"
                     /note="ORFID:SA1294"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB42554.1"
                     /db_xref="GI:13701260"
                     /translation="MTINTFWQYTLYQRSWLMMLLICNILGMIYGYIWYGEQLSHTPW
                     HFKIFVPDSPTAILYLVISISLILIQKQNSIIEALAFVTLFKYGIWAVIMNILFIIEQ
                     GDITVNGLVLMFSHSIMAVQAIYFYPRFKRSMIGISVAMIWVFLNDYIDYFHLQFPYY
                     DFITTHVWQIGVLSCCLSVFGLLLYIELNKLLKCK"
     gene            complement(1578. .2153)
                     /gene="SA1295"
     CDS             complement(1578. .2153)
                     /gene="SA1295"
                     /note="ORFID:SA1295"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB42555.1"
                     /db_xref="GI:13701261"
                     /translation="MSETLNQIKESFIEYLLFQYRFKSRIAVWVLNYIKVNEAKLANI
                     HFVDTKINHHETLEIAEVGSHASAIQFTKRNIKLMNTNEIFDYIANHNCAFDIQIHFA
                     NVSKREQRLDDLIVAQLTESPSYQTYLHDLNSMAIDRHKHALLIDYLLHNIDLSLQMN
                     EKQRFYQLTQILNTLKLVNKHNQFEDLADDD"
     gene            complement(2167. .3411)
                     /gene="SA1296"
     CDS             complement(2167. .3411)
                     /gene="SA1296"
                     /note="ORFID:SA1296"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB42556.1"
                     /db_xref="GI:13701262"
                     /translation="MEDIYKLIDDINLQKLENLDSRVNEAITTDNDDALFILGETLYN
                     FGLMPQGLEVFRVLYHKYPDESELLIYFIEGLMSENQTDEALEYLSYVEPSPEKLMLE
                     ADLYQQINMMEVAIDKLQEALELEPNDPIIHFALAEMLYYDGQYLRATSEYETVLETG
                     EYQVNGVNLFSRMADCSLQSGNYSDAIRLYDEINEDEMTSEDYLKKAISYDKNDITQE
                     AIKIMTTLLSKDPDYIQGYLYLQSLYENEKNYPDAIETGKEGLRLSQFYKELMYTTGC
                     LEIEHGDANEGVLLLKQALEVDNAYQEPLLILSDLYRNEEDYEAIIELLTYVDEEDLD
                     PTFMWHLAFAYGQEERDKEAQHFFELAYPTMKTNIDFMSDYYFYLIEIGQKEQAITIL
                     NQLLELEPSNENWHDELSRLQS"
     gene            complement(3418. .4716)
                     /gene="aroA"
     CDS             complement(3418. .4716)
                     /gene="aroA"
                     /note="ORFID:SA1297"
                     /codon_start=1
                     /transl_table=11
                     /product="3-phosphoshikimate 1-carboxyvinyltransferase"
                     /protein_id="BAB42557.1"
                     /db_xref="GI:13701263"
                     /translation="MVSEQIIDISGPLKGEIEVPGDKSMTHRAIMLASLAEGTSNIYK
                     PLLGEDCRRTMDIFRLLGVDIKEDEDKLVVNSPGYKAFKTPHQVLYTGNSGTTTRLLA
                     GLLSGLGIESVLSGDVSIGKRPMDRVLRPLKLMDANIEGIEDNYTPLIIKPSVIKGIN
                     YQMEVASAQVKSAILFASLFSNDTTVIKELDVSRNHTETMFRHFNIPIEAERLSITTT
                     PDAIQHIKPADFHVPGDISSAAFFIVAALITPESDVTIHNVGINPTRSGIIDIVEKMG
                     GNIQLFNQTTGAEPTASIRIQYTPMLQPITIEGELVPKAIDELPVIALLCTQAVGTST
                     IKDAEELKVKETNRIDTTADMLNLLGFELQPTNDGLIIHPSEFKTNATVDSLTDHRIG
                     MMLAVASLLSSEPVKIKQFDAVNVSFPGFLPKLKLLENEG"
     gene            complement(4726. .5790)
                     /gene="aroB"
     CDS             complement(4726. .5790)
                     /gene="aroB"
                     /note="ORFID:SA1298"
                     /codon_start=1
                     /transl_table=11
                     /product="3-dehydroquinate synthase"
                     /protein_id="BAB42558.1"
                     /db_xref="GI:13701264"
                     /translation="MKLQTTYPSNNYPIFVEHGAIDHISTYIDQFDQSFILIDEHVNQ
                     YFADKFDDILSYENVHKVIIPAGEKTKTFEQYQETLEYILSHHVTRNTAIIAVGGGAT
                     GDFAGFVAATLLRGVHFIQVPTTILAHDSSVGGKVGINSKQGKNLIGAFYRPTAVIYD
                     LDFLKTLPFEQILSGYAEVYKHALLNGESTTQEIEQHFKDREILQSLNGMDKYIAKGI
                     ETKLDIVVADEKEQGVRKFLNLGHTFGHAVEYNHKIAHGHAVMIGIIYQFIVANILFN
                     SNHDIQHYINYLTKLGYPLETITDIDFETIYQYMLSDKKNDKQGVQMVLIKHFGDIVV
                     QHIDQTTLQHACEQLKTYFK"
     gene            complement(5816. .6982)
                     /gene="aroC"
     CDS             complement(5816. .6982)
                     /gene="aroC"
                     /note="ORFID:SA1299"
                     /codon_start=1
                     /transl_table=11
                     /product="chorismate synthase"
                     /protein_id="BAB42559.1"
                     /db_xref="GI:13701265"
                     /translation="MRYLTSGESHGPQLTVIVEGIPANLEIKVEDINKEMFKRQGGYG
                     RGRRMQIEKDTVEIVSGVRNGYTLGSPITMVVTNDDFTHWRKIMGAAPISEEERENMK
                     RTITKPRPGHADLVGGMKYNHRDLRNVLERSSARETAARVAVGALCKVLLQQLDIDIY
                     SRVVEIGGIKDKDFYDSETFKANLDRNDVRVIDDSIAQAMRDKIDEAKNEGDSIGGVV
                     QVVVENMPVGVGSYVHYDRKLDGKIAQGVVSINAFKGVSFGEGFKAAEKPGSEIQDEI
                     LYNSEIGYYRGSNHLGGLEGGMSNGMPIIVNGVMKPIPTLYKPLNSVDINTKEDFKAT
                     IERSDSCAVPAASIVCEHVVAFEIAKALLEEFQSNHIEQLQQQIADRRQLNVEF"
     gene            complement(7368. .7568)
                     /gene="SA1300"
     CDS             complement(7368. .7568)
                     /gene="SA1300"
                     /note="ORFID:SA1300"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAB42560.1"
                     /db_xref="GI:13701266"
                     /translation="MKNIISIILGILMFLKLMELLYGAIFLDKPLNPITKIIFILTLI
                     YIFYVLVKELIIFLKSKYNKSA"
```

```
     gene            complement(7597. .7722)
                     /gene="SAS045"
     CDS             complement(7597. .7722)
                     /gene="SAS045"
                     /note="ORFID:SAS045"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAB42561.1"
                     /db_xref="GI:13701267"
                     /translation="MWLVIMIIHNKARFLCCSKLLEKFYSFLVTRIKKRGIFFIK"
     gene            complement(7777. .8226)
                     /gene="ndk"
     CDS             complement(7777. .8226)
                     /gene="ndk"
                     /note="ORFID:SA1301"
                     /codon_start=1
                     /transl_table=11
                     /product="nucleoside diphosphate kinase"
                     /protein_id="BAB42562.1"
                     /db_xref="GI:13701268"
                     /translation="MERTFLMIKPDAVQRNLIGEVISRIERKGLKLVGGKLMQVPMEL
                     AETHYGEHQGKPFYNDLISFITSAPVFAMVVEGEDAVNVSRHIIGSTNPSEASPGSIR
                     GDLGLTVGRNIIHGSDSLESAEREINLWFNENEITSYASPRDAWLYE"
     gene            complement(8318. .9277)
                     /gene="gerCC"
     CDS             complement(8318. .9277)
                     /gene="gerCC"
                     /note="ORFID:SA1302"
                     /codon_start=1
                     /transl_table=11
                     /product="heptaprenyl diphosphate syntase component II"
                     /protein_id="BAB42563.1"
                     /db_xref="GI:13701269"
                     /translation="MAKLNMNNEIKKVEQRLEKAIKSKDSVLEQASLHLLSSGGKRVR
                     PAFVILSSQFGKDEQTSEQTYQVAVALELIHMATLVHDDVIDKSDKRRGKLTISKKWD

Alignment Scores:
Pred. No.:              1.06e-126       Length:         301550
Score:                  132.00          Matches:        285
Percent Similarity:     98.62%          Conservative:   0
Best Local Similarity:  98.62%          Mismatches:     2
Query Match:            45.99%          Indels:         4
DB:                     1               Gaps:           0

US-09-925-637-64 (1-287) x AP003134 (1-301550)

Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        64130 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 64071

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64070 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGCATGTTGTATTCATTAAATGCT 64011

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64010 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 63951

Qy         61 TyrGluLeuGlyMet-LysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIl 80
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db      63950 TATGAGTTAGGTGT-GAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTAT 63892

Qy         80 eHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHi 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63891 TCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCA 63832

Qy        100 sLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPh 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63831 TAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATT 63772

Qy        120 eGluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnAr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63771 TGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACG 63712

Qy        140 gLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSe 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63711 GCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAG 63652

Qy        160 rGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAl 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63651 CGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGC 63592

Qy        180 aLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThr-ThrL 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||    ||||
Db      63591 ATTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGC-AACTA 63533

Qy        200 ysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspL 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63532 AAGAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATT 63473

Qy        220 euLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63472 TATTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAA 63413

Qy        240 snAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrT 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63412 ATAATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTT 63353

Qy        260 yrHisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysH 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63352 ATCATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAAC 63293

Qy        280 isLeuLeuGluIleValAspLeu 287
              |||||||||||||||||||||||
Db      63292 ACTTATTAGAAATCGTTGATTTA 63270

RESULT 3
AP003362/c
LOCUS       AP003362              346900 bp    DNA     linear   BCT 07-FEB-2002
DEFINITION  Staphylococcus aureus subsp. aureus Mu50 genomic DNA, complete
            sequence, section 5/9.
ACCESSION   AP003362 BA000017
VERSION     AP003362.2  GI:14247083
KEYWORDS    .
SOURCE      Staphylococcus aureus subsp. aureus Mu50 (sub_species:aureus Mu50,
            strain:Mu50) DNA.
  ORGANISM  Staphylococcus aureus subsp. aureus Mu50
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
            Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M.,
            Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
            Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
            Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
            Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
            Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
  TITLE     Whole genome sequencing of meticillin-resistant Staphylococcus
            aureus
  JOURNAL   Lancet 357 (9264), 1225-1240 (2001)
  MEDLINE   21311952
REFERENCE   2  (bases 1 to 346900)
  AUTHORS   Ohta,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Toshiko Ohta, University of Tsukuba College
            of Medical Technology and Nursing, Department of Medical
            Technology; 1-1-1 Ten-noodai, Tsukuba, Ibaraki 305-8577, Japan
            (E-mail:tohta@sakura.cc.tsukuba.ac.jp, Tel:81-298-53-3454,
            Fax:81-298-53-3454)
COMMENT     On May 29, 2001 this sequence version replaced gi:13875626.
FEATURES             Location/Qualifiers
     source          1. .346900
                     /organism="Staphylococcus aureus subsp. aureus Mu50"
                     /strain="Mu50"
                     /sub_species="aureus Mu50"
                     /db_xref="taxon:158878"
```

```
gene            12. .263
                /gene="SAV1313"
CDS             12. .263
                /gene="SAV1313"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB57475.1"
                /db_xref="GI:14247084"
                /translation="MLTLIKLKEDEQVIIYEYIPEDDISNGKGSVTFNKKDAEVIDFS
                LSEIENEEYFMLYRNKSFSVVRDFIEKQEFPENYKIAWY"
gene            987. .1181
                /gene="SAV1314"
CDS             987. .1181
                /gene="SAV1314"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB57476.1"
                /db_xref="GI:14247085"
                /translation="MNLQSYIGKLVKLTLTNNKILIGKVIDFDDKVDNFDGYNSIEID
                TGRITYDISENKIKTIVLQK"
gene            complement(1725. .2750)
                /gene="SAV1315"
CDS             complement(1725. .2750)
                /gene="SAV1315"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="BAB57477.1"
                /db_xref="GI:14247086"
                /translation="MISFENDYLEGAHEKVLKRLVDTNLVQASGYGFDQFTAQAIEKI
                KDTIDCPNATIRFLVGGTQTNQVVINSMLESYEGVISADTGHVAVHEGGAIEYSGHKV
                ITIPSKEGKVSASDVETYMETFKSDFKKDHMVFPGMVYISHPTEYGTLYSKSELEELC
                KVCKQYQLPLFMDGARLGYGLMSDQSDMTIKDIAKYCDVFYIGGTKIGALCGEAIVFT
                KNNEPKQFTTRIKHHGALLAKGRLTGIQFLELFTDNLYFNISRHAIEMANKMKDGFKN
                KGYRLYFDSPTNQQFFILSNEKIAELEQKVKFAVWEKYDDQHRVVRFATSWATTEENL
                NKLLELI"
gene            3020. .3217
                /gene="SAV1316"
CDS             3020. .3217
                /gene="SAV1316"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB57478.1"
                /db_xref="GI:14247087"
                /translation="MEQIKLKTFTAETLELLEKNINAFLSSEEATNLKLVNITIKEIE
                ERTFPNNEEEFNAILTLSVNK"
gene            3238. .4755
                /gene="SAV1317"
CDS             3238. .4755
                /gene="SAV1317"
                /codon_start=1
                /transl_table=11
                /product="cardiolipin synthetase homolog"
                /protein_id="BAB57479.1"
                /db_xref="GI:14247088"
                /translation="MFFVPFFIGEKTMRYTFSNDLGTLFTIILAIGFIINLVLAFIII
                FLERNRRTASSTWAWLFVLFVLPLIGFILYLFFGRTVSARKLNKNNGNVLTDFDGLLK
                QQIESFDKGNYGTDNKQVQKHHDLVSMLLMDQDGFLTENNKVDHFIDGNDLYDQVLKD
                IKNAKEYIHLEYYTFALDGLGKRILHALEEKLKQGLEVKILYDDVGSKNVKMANFDHF
                KSLGGEVEAFFASKLPLLNFRMNNRNHRKIIVIDGQLGYVGGFNIGDEYLGLGKLGYW
                RDTHLRIQGDAVDALQLRFILDWNSQAHRPQFEYDVKYFPKKNGPLGNSPIQIAASGP
                ASDWHQIEYGYTKMIMSAKKSVYLQSPYFIPDNSYINAIKIAAKSGVDVHLMIPCKPD
                HPLVYWATFSNASDLLSSGVKIYTYENGFIHSKMCLIDDEIVSVGTANMDFRSFELNF
                EVNAFVYDENLAKDLRVAYEHDITKSKQLTKESYANRPLSVKFKESLAKLVSPIL"
gene            4939. .5436
                /gene="SAV1318"
CDS             4939. .5436
                /gene="SAV1318"
                /note="ATP-binding protein"
                /codon_start=1
                /transl_table=11
                /product="ABC transporter homolog"
                /protein_id="BAB57480.1"
                /db_xref="GI:14247089"
                /translation="MIQISNINKSFNKRCVLKNISFDIEQGKCIALIGKNGAGKSTLI
                DILIGNVNANSGEIFDKDKLLQSENRSIMFQKTMFPDQLKVIEIINLYQSFYENPLPL
                EEIIELTKFDSSQLNQFVNKLSGGQQRLLDFVLSLIGQPQLILLDEPTSTMDIEIREY
                FWSII"
gene            5461. .5838
                /gene="SAV1319"
CDS             5461. .5838
                /gene="SAV1319"
                /note="ATP-binding protein"
                /codon_start=1
                /transl_table=11
                /product="ABC transporter homolog"
                /protein_id="BAB57481.1"
                /db_xref="GI:14247090"
                /translation="MLYTSHYIEEVERMSDKIILIENGEIILNDSTSHIRTNQQSQIT
                LSDEYIRKLKLDKDDLVIQKNHNGTIKIITSNVNDTILYLQQLHINLDDIEIQKVSIV
                DSYFNNKKQRGSNYDTKLLENRI"
gene            5807. .6538
                /gene="SAV1320"
CDS             5807. .6538
                /gene="SAV1320"
                /note="similar to ABC transporter integral membrane
                protein"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB57482.1"
                /db_xref="GI:14247091"
                /translation="MILSYLKIEFKVIMRKKTTLILSILFPVIFYILFTSILELPEDV
                KPKFYKEYMYSMTVYSLLSFSLLTFPLDIINEKQNEWRQRLMVTPFTFTSYYISKVVK
                TMLQFAIAILVIFMVGHFYKGVAMSAVQWLESGIFLWLGASLLITFGILFSLLNDIQK
                TSALANIVTIGLAVLGGLWFPINTFPNWLQHVAHVLPSYHLRKLGVDIASNHHINLIS
                FAIILLYVLGSIIAVYCISHFKRAE"
gene            6542. .7633
                /gene="SAV1321"
CDS             6542. .7633
                /gene="SAV1321"
                /note="similar to two-component sensor histidine kinase"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB57483.1"
                /db_xref="GI:14247092"
                /translation="MKFLKDTSIAEISSILYLIFPIAGIFFNEVYGPKWLYIISVIVF
                SLSYLILVIVNNRLNTLMFYILLIIHYFIICYFVFSVHPMLSLFFFYSAFAVPFTFKN
                NVKKTATNLFILTMIICTIITYLLYNNYFVAMMVYYVVISLIMLDNFKKMKNREYQKE
                IAEKNRHINTLIAEQERHRIGQDLHDTLGHVFASLSLKSELAYKLIDTDVEKVKAELL
                AINKLSRESLNKVREIIDDVKLPSFIEEIDSIRKVLKDADIDFTFENKELAQVLSPTK
                QSMLVMITREAINNVIKHANASKVHGKLKTVNNHKLLLMIEDDGKGIDSDCEVKSISQ
                RVQHLNGTLAVDSTNGTKIIIEISTGGIA"
gene            7630. .8232
                /gene="SAV1322"
CDS             7630. .8232
                /gene="SAV1322"
                /note="similar to two-component response regulator"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB57484.1"
                /db_xref="GI:14247093"
                /translation="MTSLIIAEDQNMLRQAMVQLIKLHGDFEILADTDNGLDAMKLIE
                EYNPNVVILDIEMPGMTGLEVLAEIRKKHLNIKVIIVTTFKRPGYFEKAVVNDVDAYV
                LKERSIEELVETINKVNNGEKEYSATLMTSFFVDKNPLTPKEQIVLREIGNGLSSKEI
                SEKLFLTDGTVRNYTSVIIDKLFADNRFDAWKKANEKGWI"
gene            complement(8345. .8533)
                /gene="SAV1323"
CDS             complement(8345. .8533)
                /gene="SAV1323"
                /codon_start=1
```

```
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAB57485.1"
                     /db_xref="GI:14247094"
                     /translation="MNISNENTVTLIAVIIAIIIGIFLQIFFKLPLIVTAVLSILLGI
                     FVGFIVYLIVSFYNKRKN"
     gene            8672. .9205
                     /gene="nuc"
     CDS             8672. .9205
                     /gene="nuc"
                     /note="SAV1324"
                     /codon_start=1
                     /transl_table=11
                     /product="thermonuclease"
                     /protein_id="BAB57486.1"
                     /db_xref="GI:14247095"
                     /translation="MKSNKSLAMIVVAIIIVGVLAFQFMNHTGPFKKGTNHETVQDLN

Alignment Scores:
Pred. No.:              1.2e-126      Length:         346900
Score:                  132.00        Matches:        285
Percent Similarity:     98.62%        Conservative:   0
Best Local Similarity:  98.62%        Mismatches:     2
Query Match:            45.99%        Indels:         4
DB:                     1             Gaps:           0

US-09-925-637-64 (1-287) x AP003362 (1-346900)

Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251159 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 251100

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251099 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGCATGTTGTATTCATTAAATGCT 251040

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251039 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 250980

Qy         61 TyrGluLeuGlyMet-LysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIl 80
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||
Db     250979 TATGAGTTAGGTGT-GAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTAT 250921

Qy         80 eHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHi 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250920 TCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCA 250861

Qy        100 sLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPh 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250860 TAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATT 250801

Qy        120 eGluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnAr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250800 TGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACG 250741

Qy        140 gLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSe 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250740 GCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAG 250681

Qy        160 rGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAl 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250680 CGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGC 250621

Qy        180 aLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThr-ThrL 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||    ||||
Db     250620 ATTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGC-AACTA 250562

Qy        200 ysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspL 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250561 AAGAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATT 250502

Qy        220 euLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250501 TATTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAA 250442

Qy        240 snAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrT 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250441 ATAATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTT 250382

Qy        260 yrHisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysH 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250381 ATCATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAAC 250322

Qy        280 isLeuLeuGluIleValAspLeu 287
              |||||||||||||||||||||||
Db     250321 ACTTATTAGAAATCGTTGATTTA 250299

RESULT 4
AR106466
LOCUS       AR106466                882 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 1 from patent US 6107058.
ACCESSION   AR106466
VERSION     AR106466.1  GI:12820996
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 882)
  AUTHORS   Gwynn,M. and Wilding,E.Imogen.
  TITLE     ispA from Staphylococcus aureus
  JOURNAL   Patent: US 6107058-A 1 22-AUG-2000;
FEATURES             Location/Qualifiers
     source          1. .882
                     /organism="unknown"
BASE COUNT      328 a    118 c    186 g    250 t
ORIGIN

Alignment Scores:
Pred. No.:              1.91e-97      Length:         882
Score:                  102.00        Matches:        284
Percent Similarity:     97.93%        Conservative:   0
Best Local Similarity:  97.93%        Mismatches:     3
Query Match:            35.54%        Indels:         6
DB:                     6             Gaps:           0

US-09-925-637-64 (1-287) x AR106466 (1-882)

Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGCATGTTGTATTCATTAAATGCT 120

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy         61 TyrGluLeuGlyMet-LysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIl 80
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTGT-GAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTAT 239

Qy         80 eHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGly-LysLeuThrAsnH 100
              ||||||||||||||||||||||||||||||||||||||||||||    ||||||||||||
Db        240 TCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAAG-GAAATTAACAAATC 298

Qy        100 isLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCAT 358

Qy        120 heGluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnA 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 TTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAAC 418
```

```
Qy        140 rgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnS 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 GGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAA 478

Qy        160 erGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        479 GCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAG 538

Qy        180 laLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThr-Thr 199
              |||||||||||||||||||||||||||||||||||||||||||||||||||||    |||
Db        539 CATTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGC-AACT 597

Qy        200 LysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAsp 219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 AAAGAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGAT 657

Qy        220 LeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGlu 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 TTATTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAA 717

Qy        240 AsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThr 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 AATAATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACT 777

Qy        260 TyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLys 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 TATCATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAA 837

Qy        280 HisLeuLeuGluIleValAspLeu 287
              |||||||||||||||||||||||
Db        838 CACTTATTAGAAATCGTTGATTTA 861
```

RESULT 5
AX141501

LOCUS	AX141501	783 bp	DNA	linear	PAT 31-MAY-2001
DEFINITION	Sequence 223 from Patent WO0134809.				
ACCESSION	AX141501				
VERSION	AX141501.1 GI:14281622				
KEYWORDS	.				
SOURCE	synthetic construct.				
ORGANISM	synthetic construct artificial sequences.				
REFERENCE	1 (bases 1 to 783)				
AUTHORS	Kimmerly,W.J.				
TITLE	Staphylococcus epidermidis nucleic acids and proteins				
JOURNAL	Patent: WO 0134809-A 223 17-MAY-2001; GLAXO GROUP LIMITED (GB)				

FEATURES	Location/Qualifiers
source	1. .783 /organism="synthetic construct" /db_xref="taxon:32630" /note="synthetic nucleic acid sequence"

BASE COUNT 297 a 93 c 147 g 246 t
ORIGIN

Alignment Scores:

Pred. No.:	3.62e-29	Length:	783
Score:	37.00	Matches:	37
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.89%	Indels:	0
DB:	6	Gaps:	0

US-09-925-637-64 (1-287) x AX141501 (1-783)

```
Qy         70 AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsn 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        109 GCATTGGAAATGATTCATACTTATTCTTTAATTCATGATGATTTACCAGCAATGGATAAT 168

Qy         90 AspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrp 106
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 GACGATTACCGTAGAGGAAAATTAACAAATCATAAAGTTTATGGTGAATGG 219
```

RESULT 6
AF270104

LOCUS	AF270104	3444 bp	DNA	linear	BCT 01-AUG-2000
DEFINITION	Staphylococcus epidermidis strain SR1 clone step.1049b08 genomic sequence.				
ACCESSION	AF270104				
VERSION	AF270104.1 GI:9624010				
KEYWORDS	.				
SOURCE	Staphylococcus epidermidis.				
ORGANISM	Staphylococcus epidermidis Bacteria; Firmicutes; Bacillales; Staphylococcus.				
REFERENCE	1 (bases 1 to 3444)				
AUTHORS	Kimmerly,W.J., Taylor,J.David., Nelsen,A.J., Godlevski,M.M., Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I., Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S., Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and Furdon,P.J.				
TITLE	Transposon-mediated sequencing of the Staphylococcus epidermidis genome				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 3444)				
AUTHORS	Taylor,J.David., Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M., Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I., Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S., Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and Furdon,P.J.				
TITLE	Direct Submission				
JOURNAL	Submitted (22-MAY-2000) Departments of Genomic Sciences and Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore Drive, Research Triangle Park, North Carolina 27709-3398, USA				

FEATURES	Location/Qualifiers
source	1. .3444 /organism="Staphylococcus epidermidis" /strain="SR1" /db_xref="taxon:1282" /clone="step.1049b08"

BASE COUNT 1365 a 451 c 571 g 1057 t
ORIGIN

Alignment Scores:

Pred. No.:	1.37e-28	Length:	3444
Score:	37.00	Matches:	37
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.89%	Indels:	0
DB:	1	Gaps:	0

US-09-925-637-64 (1-287) x AF270104 (1-3444)

```
Qy         70 AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsn 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        973 GCATTGGAAATGATTCATACTTATTCTTTAATTCATGATGATTTACCAGCAATGGATAAT 1032

Qy         90 AspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrp 106
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1033 GACGATTACCGTAGAGGAAAATTAACAAATCATAAAGTTTATGGTGAATGG 1083
```

RESULT 7
AX145422

LOCUS	AX145422	3444 bp	DNA	linear	PAT 31-MAY-2001
DEFINITION	Sequence 4144 from Patent WO0134809.				
ACCESSION	AX145422				
VERSION	AX145422.1 GI:14283987				
KEYWORDS	.				
SOURCE	synthetic construct.				
ORGANISM	synthetic construct artificial sequences.				
REFERENCE	1 (bases 1 to 3444)				
AUTHORS	Kimmerly,W.J.				
TITLE	Staphylococcus epidermidis nucleic acids and proteins				
JOURNAL	Patent: WO 0134809-A 4144 17-MAY-2001; GLAXO GROUP LIMITED (GB)				

FEATURES Location/Qualifiers

```
     source          1. .3444
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic nucleic acid sequence"
BASE COUNT     1365 a    451 c    571 g   1057 t
ORIGIN

Alignment Scores:
Pred. No.:              1.37e-28      Length:         3444
Score:                  37.00         Matches:        37
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            12.89%        Indels:         0
DB:                     6             Gaps:           0

US-09-925-637-64 (1-287) x AX145422 (1-3444)

Qy         70 AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsn 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        973 GCATTGGAAATGATTCATACTTATTCTTTAATTCATGATGATTTACCAGCAATGGATAAT 1032

Qy         90 AspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrp 106
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1033 GACGATTACCGTAGAGGAAAATTAACAAATCATAAAGTTTATGGTGAATGG 1083

RESULT 8
AF269958/c
LOCUS       AF269958               3454 bp    DNA     linear   BCT 01-AUG-2000
DEFINITION  Staphylococcus epidermidis strain SR1 clone step.1035g12 genomic
            sequence.
ACCESSION   AF269958
VERSION     AF269958.1  GI:9623860
KEYWORDS    .
SOURCE      Staphylococcus epidermidis.
  ORGANISM  Staphylococcus epidermidis
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1  (bases 1 to 3454)
  AUTHORS   Kimmerly,W.J., Taylor,J.David., Nelsen,A.J., Godlevski,M.M.,
            Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
            Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,
            Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
            Furdon,P.J.
  TITLE     Transposon-mediated sequencing of the Staphylococcus epidermidis
            genome
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3454)
  AUTHORS   Taylor,J.David., Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M.,
            Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
            Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,
            Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
            Furdon,P.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAY-2000) Departments of Genomic Sciences and
            Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
            Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES              Location/Qualifiers
     source           1. .3454
                      /organism="Staphylococcus epidermidis"
                      /strain="SR1"
                      /db_xref="taxon:1282"
                      /clone="step.1035g12"
BASE COUNT     1005 a    574 c    493 g   1382 t
ORIGIN

Alignment Scores:
Pred. No.:              1.38e-28      Length:         3454
Score:                  37.00         Matches:        37
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            12.89%        Indels:         0
DB:                     1             Gaps:           0

US-09-925-637-64 (1-287) x AF269958 (1-3454)

Qy         70 AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsn 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 GCATTGGAAATGATTCATACTTATTCTTTAATTCATGATGATTTACCAGCAATGGATAAT 91

Qy         90 AspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrp 106
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 GACGATTACCGTAGAGGAAAATTAACAAATCATAAAGTTTATGGTGAATGG 40

RESULT 9
AX145276/c
LOCUS       AX145276               3454 bp    DNA     linear   PAT 31-MAY-2001
DEFINITION  Sequence 3998 from Patent WO0134809.
ACCESSION   AX145276
VERSION     AX145276.1  GI:14283841
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 3454)
  AUTHORS   Kimmerly,W.J.
  TITLE     Staphylococcus epidermidis nucleic acids and proteins
  JOURNAL   Patent: WO 0134809-A 3998 17-MAY-2001;
            GLAXO GROUP LIMITED (GB)
FEATURES              Location/Qualifiers
     source           1. .3454
                      /organism="synthetic construct"
                      /db_xref="taxon:32630"
                      /note="synthetic nucleic acid sequence"
BASE COUNT     1005 a    574 c    493 g   1382 t
ORIGIN

Alignment Scores:
Pred. No.:              1.38e-28      Length:         3454
Score:                  37.00         Matches:        37
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            12.89%        Indels:         0
DB:                     6             Gaps:           0

US-09-925-637-64 (1-287) x AX145276 (1-3454)

Qy         70 AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsn 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 GCATTGGAAATGATTCATACTTATTCTTTAATTCATGATGATTTACCAGCAATGGATAAT 91

Qy         90 AspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrp 106
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 GACGATTACCGTAGAGGAAAATTAACAAATCATAAAGTTTATGGTGAATGG 40

RESULT 10
AF269889/c
LOCUS       AF269889               4045 bp    DNA     linear   BCT 01-AUG-2000
DEFINITION  Staphylococcus epidermidis strain SR1 clone step.1028f08 genomic
            sequence.
ACCESSION   AF269889
VERSION     AF269889.1  GI:9623789
KEYWORDS    .
SOURCE      Staphylococcus epidermidis.
  ORGANISM  Staphylococcus epidermidis
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1  (bases 1 to 4045)
  AUTHORS   Kimmerly,W.J., Taylor,J.David., Nelsen,A.J., Godlevski,M.M.,
            Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
            Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,
            Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
            Furdon,P.J.
  TITLE     Transposon-mediated sequencing of the Staphylococcus epidermidis
            genome
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4045)
  AUTHORS   Taylor,J.David., Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M.,
            Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
```

```
          Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,
          Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
          Furdon,P.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAY-2000) Departments of Genomic Sciences and
          Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
          Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES             Location/Qualifiers
     source          1. .4045
                     /organism="Staphylococcus epidermidis"
                     /strain="SR1"
                     /db_xref="taxon:1282"
                     /clone="step.1028f08"
BASE COUNT     1249 a    691 c    536 g   1569 t
ORIGIN

Alignment Scores:
Pred. No.:                1.59e-28       Length:        4045
Score:                    37.00          Matches:       37
Percent Similarity:       100.00%        Conservative:  0
Best Local Similarity:    100.00%        Mismatches:    0
Query Match:              12.89%         Indels:        0
DB:                       1              Gaps:          0

US-09-925-637-64 (1-287) x AF269889 (1-4045)

Qy          70 AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsn 89
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2106 GCATTGGAAATGATTCATACTTATTCTTTAATTCATGATGATTTACCAGCAATGGATAAT 2047

Qy          90 AspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrp 106
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2046 GACGATTACCGTAGAGGAAAATTAACAAATCATAAAGTTTATGGTGAATGG 1996

RESULT 11
AX145207/c
LOCUS       AX145207               4045 bp    DNA     linear   PAT 31-MAY-2001
DEFINITION  Sequence 3929 from Patent WO0134809.
ACCESSION   AX145207
VERSION     AX145207.1  GI:14283772
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4045)
  AUTHORS   Kimmerly,W.J.
  TITLE     Staphylococcus epidermidis nucleic acids and proteins
  JOURNAL   Patent: WO 0134809-A 3929 17-MAY-2001;
            GLAXO GROUP LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .4045
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic nucleic acid sequence"
BASE COUNT     1249 a    691 c    536 g   1569 t
ORIGIN

Alignment Scores:
Pred. No.:                1.59e-28       Length:        4045
Score:                    37.00          Matches:       37
Percent Similarity:       100.00%        Conservative:  0
Best Local Similarity:    100.00%        Mismatches:    0
Query Match:              12.89%         Indels:        0
DB:                       6              Gaps:          0

US-09-925-637-64 (1-287) x AX145207 (1-4045)

Qy          70 AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsn 89
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2106 GCATTGGAAATGATTCATACTTATTCTTTAATTCATGATGATTTACCAGCAATGGATAAT 2047

Qy          90 AspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrp 106
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2046 GACGATTACCGTAGAGGAAAATTAACAAATCATAAAGTTTATGGTGAATGG 1996

RESULT 12
AR097529
LOCUS       AR097529               486 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 28 from patent US 6071733.
ACCESSION   AR097529
VERSION     AR097529.1  GI:12806259
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 486)
  AUTHORS   Muramatsu,M., Koike,A., Ogura,K., Koyama,T., Shimizu,N. and Cho,Y.
  TITLE     Method of making recombinant enzyme
  JOURNAL   Patent: US 6071733-A 28 06-JUN-2000;
FEATURES             Location/Qualifiers
     source          1. .486
                     /organism="unknown"
BASE COUNT     148 a     85 c     98 g    149 t      6 others
ORIGIN

Alignment Scores:
Pred. No.:                3.78e-25       Length:        486
Score:                    33.00          Matches:       33
Percent Similarity:       100.00%        Conservative:  0
Best Local Similarity:    100.00%        Mismatches:    0
Query Match:              11.50%         Indels:        0
DB:                       6              Gaps:          0

US-09-925-637-64 (1-287) x AR097529 (1-486)

Qy          69 IleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAsp 88
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          85 ATTGCGCTTGAAATGATTCATACGTATTCTTTAATTCACGATGACTTGCCGGCAATGGAT 144

Qy          89 AsnAspAspTyrArgArgGlyLysLeuThrAsnHisLys 101
               |||||||||||||||||||||||||||||||||||||||
Db         145 AATGATGACTATCGTCGCGGTAAACTCACGAATCATAAG 183

RESULT 13
AR142282
LOCUS       AR142282               486 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 28 from patent US 6174715.
ACCESSION   AR142282
VERSION     AR142282.1  GI:15102582
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 486)
  AUTHORS   Muramatsu,M., Koike,A., Ogura,K., Koyama,T., Shimizu,N. and Cho,Y.
  TITLE     Prenyl diphosphate synthetase genes
  JOURNAL   Patent: US 6174715-A 28 16-JAN-2001;
FEATURES             Location/Qualifiers
     source          1. .486
                     /organism="unknown"
BASE COUNT     148 a     85 c     98 g    149 t      6 others
ORIGIN

Alignment Scores:
Pred. No.:                3.78e-25       Length:        486
Score:                    33.00          Matches:       33
Percent Similarity:       100.00%        Conservative:  0
Best Local Similarity:    100.00%        Mismatches:    0
Query Match:              11.50%         Indels:        0
DB:                       6              Gaps:          0

US-09-925-637-64 (1-287) x AR142282 (1-486)

Qy          69 IleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAsp 88
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          85 ATTGCGCTTGAAATGATTCATACGTATTCTTTAATTCACGATGACTTGCCGGCAATGGAT 144
```

```
Qy          89 AsnAspAspTyrArgArgGlyLysLeuThrAsnHisLys 101
               ||||||||||||||||||||||||||||||||||||||||
Db         145 AATGATGACTATCGTCGCGGTAAACTCACGAATCATAAG 183

RESULT 14
E15319
LOCUS       E15319                   486 bp    DNA     linear   PAT 28-JUL-1999
DEFINITION  Micrococcus luteus FPS gene fragment.
ACCESSION   E15319
VERSION     E15319.1  GI:5710002
KEYWORDS    JP 1998057079-A/10.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 486)
  AUTHORS   Muramatsu,M., Koike,A., Ogura,K., Furuyama,T., Shimizu,N. and
            Chiyou,G.
  TITLE     GENE FOR PRENYL DIPHOSPHORIC ACID-SYNTHESIZING ENZYME
  JOURNAL   Patent: JP 1998057079-A 10 03-MAR-1998;
            TOYOTA MOTOR CORP
COMMENT     OS   Micrococcus luteus
            PN   JP 1998057079-A/10
            PD   03-MAR-1998
            PF   13-JUN-1997 JP 1997157228
            PR   14-JUN-1996 JP 96P   154441
            PI   MURAMATSU MASAYOSHI, KOIKE AYUMI, OGURA KYOZO, PI   FURUYAMA
            TANETOSHI,
            PI   SHIMIZU NAOTO, CHIYOU GENI
            PC   C12N15/09,C07H21/04,C12N1/21,C12N9/00,(C12N15/09,C12R1:125),
            PC   (C12N15/09,
            PC   C12R1:07),(C12N1/21,C12R1:19),(C12N9/00,C12R1:19); CC
            strandedness: Double;
            CC   topology: Linear;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .486
            FT                   /organism='Microsoccus luteus'.
FEATURES             Location/Qualifiers
     source          1. .486
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      148 a     85 c     98 g    149 t      6 others
ORIGIN

Alignment Scores:
Pred. No.:              3.78e-25      Length:         486
Score:                  33.00         Matches:        33
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            11.50%        Indels:         0
DB:                     6             Gaps:           0

US-09-925-637-64 (1-287) x E15319 (1-486)

Qy          69 IleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAsp 88
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          85 ATTGCGCTTGAAATGATTCATACGTATTCTTTAATTCACGATGACTTGCCGGCAATGGAT 144

Qy          89 AsnAspAspTyrArgArgGlyLysLeuThrAsnHisLys 101
               ||||||||||||||||||||||||||||||||||||||||
Db         145 AATGATGACTATCGTCGCGGTAAACTCACGAATCATAAG 183

RESULT 15
AB003187
LOCUS       AB003187                1080 bp    DNA     linear   BCT 25-MAR-1998
DEFINITION  Micrococcus luteus DNA for farnesyl diphosphate synthase, complete
            cds.
ACCESSION   AB003187
VERSION     AB003187.1  GI:2982676
KEYWORDS    farnesyl diphosphate synthase.
SOURCE      Micrococcus luteus (strain:B-P 26) DNA.
  ORGANISM  Micrococcus luteus
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Micrococcineae; Micrococcaceae; Micrococcus.
REFERENCE   1  (sites)
  AUTHORS   Shimizu,N., Koyama,T. and Ogura,K.
  TITLE     Molecular cloning, expression, and characterization of the genes
            encoding the two essential protein components of Micrococcus luteus
            B-P 26 hexaprenyl diphosphate synthase
  JOURNAL   J. Bacteriol. 180 (6), 1578-1581 (1998)
  MEDLINE   98175686
REFERENCE   2  (bases 1 to 1080)
  AUTHORS   koyama,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-APR-1997) Tanetoshi koyama, Tohoku University,
            Institute for Chemical Reaction Science; Katahira 2-1-1, Aoba-ku,
            Sendai, Miyagi 980-77, Japan (E-mail:koyama@icrs.tohoku.ac.jp,
            Tel:+81-22-217-5621, Fax:+81-22-217-5620)
FEATURES             Location/Qualifiers
     source          1. .1080
                     /organism="Micrococcus luteus"
                     /strain="B-P 26"
                     /db_xref="taxon:1270"
     gene            105. .980
                     /gene="fps"
     CDS             105. .980
                     /gene="fps"
                     /codon_start=1
                     /transl_table=11
                     /product="farnesyl diphosphate synthase"
                     /protein_id="BAA25265.1"
                     /db_xref="GI:2982677"
                     /translation="MLQEKLTMNRDFLNLINESLLNKYHPAQSRLHEAINYSLSAGGK
                     RIRPLLVLTTLDSLGGNAHDGLPFGIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHK
                     RFDEATAILAGDALLTDAFQCILNTQLNAEIKLSLINLLSTASGSNGMVYGQMLDMQG
                     EHKTLTLNELERIHIHKTGELIRAAIVSAGIIMNFNDAQIEQLNIIGKNVGLMFQIKD
                     DILDVEGSFENIGKTVGSDLNNDKSTYVSLLGLEASKQLLNDKLTETYDALKTLQPIN
                     DNLKTLITYIVERNK"
BASE COUNT      377 a    161 c    198 g    344 t
ORIGIN

Alignment Scores:
Pred. No.:              7.75e-25      Length:         1080
Score:                  33.00         Matches:        33
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            11.50%        Indels:         0
DB:                     1             Gaps:           0

US-09-925-637-64 (1-287) x AB003187 (1-1080)

Qy          69 IleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAsp 88
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         312 ATTGCGCTTGAAATGATTCATACGTATTCTTTAATTCACGATGACTTGCCGGCAATGGAT 371

Qy          89 AsnAspAspTyrArgArgGlyLysLeuThrAsnHisLys 101
               ||||||||||||||||||||||||||||||||||||||||
Db         372 AATGATGACTATCGTCGCGGTAAACTCACGAATCATAAG 410


Search completed: May 30, 2003, 23:41:52
Job time : 2998 secs
```

PD 30-JUL-1997.
XX
PF 07-JAN-1997; 97EP-0100117.
XX
PR 05-JAN-1996; 96US-0009861.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
PI Rosen CA;
XX
DR WPI; 1997-374922/35.
XX
PT Polynucleotide(s) and proteins derived from Staphylococcus aureus -
PT stored on computer readable medium and used in the production of
PT anti-S.aureus vaccines
XX
PS Claim 1; Page 784-785; 3271pp; English.
XX
CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the S.aureus DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against S.aureus infection. The
CC polypeptides can also be used in a kit for the immunodetection of
CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC skin and surgical wound infections, scalded skin syndrome, toxic shock
CC syndrome, etc. Organisms transformed with the DNA sequences can be used
CC for recombinant production of the polypeptides. The new DNA sequences
CC (and their fragments) are useful as primers or probes for isolating
CC homologues of any of the S.aureus DNA sequences contained on the
CC computer readable medium.
XX
SQ Sequence 1893 BP; 552 A; 336 C; 233 G; 708 T; 64 other;

Alignment Scores:

Pred. No.:	9.02e-262	Length:	1893
Score:	269.00	Matches:	269
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	93.73%	Indels:	0
DB:	18	Gaps:	0

US-09-925-637-64 (1-287) x AAV74466 (1-1893)

```
Qy         19 ValAlaIleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeu 38
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GTTGCGATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTA 1201

Qy         39 AsnAlaGlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsn 58
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 AATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAAT 1141

Qy         59 ThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSer 78
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 ACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCA 1081

Qy         79 LeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThr 98
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACA 1021

Qy         99 AsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLys 118
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 AATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAA 961

Qy        119 AlaPheGluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeu 138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 GCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTA 901

Qy        139 GlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMet 158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 CAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATG 841

Qy        159 GlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThr 178
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 CAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACA 781

Qy        179 GlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThr 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACA 721

Qy        199 ThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAsp 218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACTAAAGAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGAT 661

Qy        219 AspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeu 238
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 GATTTATTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTT 601

Qy        239 GluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeu 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GAAAATAATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTG 541

Qy        259 ThrTyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThr 278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ACTTATCATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACA 481

Qy        279 LysHisLeuLeuGluIleValAspLeu 287
              |||||||||||||||||||||||||||
Db        480 AAACACTTATTAGAAATCGTTGATTTA 454
```

RESULT 3
AAS54848
ID AAS54848 standard; DNA; 882 BP.
XX
AC AAS54848;
XX
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus DNA for cellular proliferation protein #1160.
XX
KW Antisense; ds; prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR P-PSDB; AAU36989.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX

```
Db        181 AAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCA 240

Qy        256 GCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAG 315
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        241 GCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCATAAAGTATATGGTGAG 300

Qy        316 TGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGT 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGT 360

Qy        376 GATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGT 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGT 420

Qy        436 GGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATT 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATT 480

Qy        496 GATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTTTTGCG 555
              |||||||||||||||||||||||||||||||||||||||||||| ||||||| ||||||
Db        481 GATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCACTATTAACATTTGCG 540

Qy        556 GTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAGAAAGT 615
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        541 GTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAAGAACATTTAGAAAGT 600

Qy        616 TATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCTATGGT 675
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        601 TATAGTTATCATTTAGGTATGATGTTTCAGATTAAAGATGATTTATTAGACTGCTATGGT 660

Qy        676 GATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTACGTAC 735
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||| |||
Db        661 GATGAAGCGAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTACATAC 720

Qy        736 GTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGACGCAGCA 795
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTGAGTTTATTAGGAAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGACGCAGCA 780

Qy        796 GTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACACTTATTAGAAATCGTT 855
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GTGGATGAACTAATGCAAATTGATGAACAATTCAATACAAAACACTTATTAGAAATCGTT 840

Qy        856 GATTTA 861
              ||||||
Db        841 GATTTA 846


RESULT 5
AAV74466/c
ID   AAV74466 standard; DNA; 1893 BP.
XX
AC   AAV74466;
XX
DT   16-MAR-1999  (first entry)
XX
DE   Staphylococcus aureus contig SEQ ID #155.
XX
KW   Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW   cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW   skin infection; surgical wound infection; scalded skin syndrome;
KW   toxic shock syndrome; ds.
XX
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1261..1320
FT                   /*tag= a
FT                   /note= "these bases represent a line of missing text in
FT                         the sequence listing in the specification. They
FT                         are included to maintain the nucleotide numbering
FT                         given in the specification for this DNA sequence"
XX
PN   EP786519-A2.
XX
PD   30-JUL-1997.
XX
PF   07-JAN-1997;   97EP-0100117.
XX
PR   05-JAN-1996;   96US-0009861.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Barash SC,  Choi GH,  Dillon PJ,  Fannon MR,  Kunsch CA;
PI   Rosen CA;
XX
DR   WPI; 1997-374922/35.
XX
PT   Polynucleotide(s) and proteins derived from Staphylococcus aureus -
PT   stored on computer readable medium and used in the production of
PT   anti-S.aureus vaccines
XX
PS   Claim 1; Page 784-785; 3271pp; English.
XX
CC   This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC   of the invention. The DNA sequences are recorded on a computer readable
CC   medium, preferably selected from a floppy or hard disk, random access
CC   memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC   the S.aureus DNA sequences allows putative functions to be assigned so
CC   that protein-encoding or regulatory regions of commercial, therapeutic or
CC   industrial importance can be obtained. Specifically, sequences which are
CC   likely to encode antigens have been identified and these polypeptides can
CC   be used in a vaccine composition against S.aureus infection. The
CC   polypeptides can also be used in a kit for the immunodetection of
CC   S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC   including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC   skin and surgical wound infections, scalded skin syndrome, toxic shock
CC   syndrome, etc. Organisms transformed with the DNA sequences can be used
CC   for recombinant production of the polypeptides. The new DNA sequences
CC   (and their fragments) are useful as primers or probes for isolating
CC   homologues of any of the S.aureus DNA sequences contained on the
CC   computer readable medium.
XX
SQ   Sequence 1893 BP; 552 A; 336 C; 233 G; 708 T; 64 other;

  Query Match             93.7%;  Score 807;  DB 18;  Length 1893;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-181;
  Matches  807;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         55 GTTGCGATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTA 114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GTTGCGATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTA 12

Qy        115 AATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAAT 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 AATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAAT 1141

Qy        175 ACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCA 234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 ACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCA 1081

Qy        235 CTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACA 294
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACA 1021

Qy        295 AATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAA 354
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 AATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAA 961

Qy        355 GCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTA 414
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 GCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTA 901

Qy        415 CAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATG 474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 CAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATG 841
```

```
Qy        475 CAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACA 534
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 CAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACA 781

Qy        535 GGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACA 594
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACA 721

Qy        595 ACTAAAGAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGAT 654
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACTAAAGAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGAT 661

Qy        655 GATTTATTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTT 714
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 GATTTATTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTT 601

Qy        715 GAAAATAATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTG 774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GAAAATAATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTG 541

Qy        775 ACTTATCATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACA 834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ACTTATCATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACA 481

Qy        835 AAACACTTATTAGAAATCGTTGATTTA 861
              |||||||||||||||||||||||||||
Db        480 AAACACTTATTAGAAATCGTTGATTTA 454
```

RESULT 6
ABN90871
ID ABN90871 standard; DNA; 909 BP.
XX
AC ABN90871;
XX
DT 24-JUL-2002 (first entry)
XX
DE Staphylococcus epidermidis ORF nucleic acid sequence SEQ ID NO:334.
XX
KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW antibacterial; gene therapy; gene; ds.
XX
OS Staphylococcus epidermidis.
XX
PN US6380370-B1.
XX
PD 30-APR-2002.
XX
PF 13-AUG-1998; 98US-0134001.
XX
PR 14-AUG-1997; 97US-055779P.
PR 08-NOV-1997; 97US-064964P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Doucette-Stamm LA, Bush D;
XX
DR WPI; 2002-381255/41.
DR P-PSDB; ABP38326.
XX
PT Novel isolated nucleic acid encoding a Staphylococcus epidermidis
PT polypeptide, useful for diagnosing and treating bacterial infections -
XX
PS Disclosure; SEQ ID 334; 267pp; English.
XX
CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC antibacterial activity and can be used in gene therapy. The sequences
CC can also be used in the diagnosis and treatment of bacterial infections,
CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life
CC cycle or inhibit S. epidermidis infection.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site.
XX
SQ Sequence 909 BP; 361 A; 106 C; 165 G; 277 T; 0 other;

Query Match 52.2%; Score 449.8; DB 24; Length 909;
Best Local Similarity 70.2%; Pred. No. 6.2e-97;
Matches 604; Conservative 0; Mismatches 257; Indels 0; Gaps 0;

```
Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              |||| ||| |||| ||||||||||||||||| ||  | | ||||    | |       |
Db         28 ATGAAGAAACTACAGATGAATAAATTAATAAATATAATTAATACATCACTGAATAAGTCA 87

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||| |   |||| | |  | ||| |  ||||||||||||||   ||||||||||||||||
Db         88 ATACAATCATCACCATTAAAAACTAATTTAGAAGAAAGTATGAAATATTCATTAAATGCT 147

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              || |||||| | ||| |||||||  | ||||||  | ||  || |    || || |  ||
Db        148 GGTGGTAAAAGAATCAGACCAGTCATATTATTATTAACACTAAAAATGCTTAACAAAGAT 207

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              ||| |   |||  | || || ||  | ||| | ||||||||||||||| |||||  | |||
Db        208 TATCAACAAGGACTAAATAGTGCTTTAGCATTGGAAATGATTCATACTTATTCTTTAATT 267

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              ||||||||  |||||||  |||||||||||  |||||  ||  ||||||||||||||||||
Db        268 CATGATGATTTACCAGCAATGGATAATGACGATTACCGTAGAGGAAAATTAACAAATCAT 327

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              |||||  ||||||||  ||||  ||  ||  |  ||  ||||||||  ||||||||  ||||| |||
Db        328 AAAGTTTATGGTGAATGGAAAGCCATTCTTGCTGGTGATGCATTATTAACAAAAGCTTTT 387

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              |||  |  ||||  |  ||||   ||    |    |||   ||  |||  |||    ||  ||  ||  |
Db        388 GAATTAGTTTCTAATGATACTACCATTGAAGATAGTGTGAAAGTAAGTATTATAAAAAGA 447

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||  |||| |||||||||  |||  |  ||||||||  ||  ||  |||   |  |  ||||||  ||||
Db        448 CTTTCAAAAGCAAGTGGACATTTGGGAATGGTGGGTGGCCAAGCGCTTGATATGGAAAGT 507

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              |||||  |  |||||  |  |  ||||||||  |||   ||  ||  ||||  ||  |||||  ||
Db        508 GAAGGGAAGTCAATTCGTTTAGAAACTTTAGAATCAATTCATGAAACTAAGACAGGCGCT 567

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              |||  |||  ||||  |  ||||||   |||  |  |||  ||  |||  |  ||  ||   |  |  |
Db        568 TTACTAAATTTTTCAGTTATGGCTGCGGTAGACATTGCTCAAGTAGAACAAAATATTGCT 627

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              |  |||||||    |  ||||  ||||||||||  ||||||||  ||  ||||||||||||||||
Db        628 AAGAATTTAGATGAATTTAGTCATCATTTAGGAATGATGTTTCAAATTAAAGATGATTTA 687

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              |  ||    ||||||||||||  ||||  |  ||  ||||||||  |||||  |||  |  |  ||||
Db        688 CTGGATGTGTATGGTGATGAATCAAAACTTGGCAAAAAAGTAGGCAGTGATATAGTAAAT 747

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||||  ||  ||||  |  ||  |||||  ||  ||||||||  ||  ||  |   ||
Db        748 CATAAAAGTACTTATGTTTCTTTACTTGGAAAAGAAGGAGCAGAAGAAAAGTTAAACAAT 807

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              |||  |  |    ||  ||  |     |||  ||||||  |||  ||||   ||||      |
Db        808 CATCAATATCTTGCTATGAACTGCTTAAATCAAATTTCTGATCAATATGATACTTCTGAA 867

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||   || ||  || ||||||
Db        868 TTAAGTGATATTGTAGATTTA 888
```

RESULT 7

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 30, 2003, 21:50:41 ; Search time 276 Seconds
(without alignments)
2341.752 Million cell updates/sec

Title: US-09-925-637-64
Perfect score: 287
Sequence: 1 MTNLPMNKLIDEVNNELSVA.....ELTQIDEQFNTKHLLEIVDL 287

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4368727

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09925637/runat_23052003_174850_24102/app_query.fasta_1.455
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=oli.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09925637_@CGN_1_1_263_@runat_23052003_174850_24102 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	Query % Match	Length	DB	ID	Description
1		287	100.0	861	22	AAS00821	S. aureus HGS072 e
2	c	269	93.7	1893	18	AAV74466	Staphylococcus aur
3		197	68.6	882	23	AAS54848	Staphylococcus aur
4		192	66.9	864	23	AAS51602	Staphylococcus aur
5	c	137	47.7	413	23	AAS50245	Staphylococcus aur
6	c	112	39.0	337	23	AAS50698	Staphylococcus aur
7		102	35.5	882	21	AAA92031	Staphylococcus aur
8		37	12.9	783	22	AAH52415	S. epidermidis ope
9		37	12.9	909	24	ABN90871	Staphylococcus epi
10		37	12.9	3444	22	AAH54780	S. epidermidis gen
11	c	37	12.9	3454	22	AAH54634	S. epidermidis gen
12	c	37	12.9	4045	22	AAH54565	S. epidermidis gen
13		33	11.5	486	19	AAV15824	Probe B500 for Mic
14		27	9.4	5635	24	ABQ71010	Listeria monocytog
15		27	9.4	2944528	24	ABA03041	Listeria monocytog
16	c	19	6.6	495269	24	ABQ67195	Listeria innocua c
17		19	6.6	3011208	24	ABQ69245	Listeria innocua D
18		18	6.3	870	24	ABN66185	Streptococcus poly
19		18	6.3	882	23	AAS52896	Enterococcus faeca
20		18	6.3	7528	20	AAX12992	Enterococcus faeca
21		18	6.3	2365589	24	ABA90521	Genomic sequence o
22		15	5.2	891	24	ABK83265	High growth methan
23		15	5.2	891	24	AAD35506	Methylomonas 16a s
24		15	5.2	891	24	ABK50087	Methylomonas 16a O
25		14	4.9	811	24	ABK73724	Bacillus lichenifo
26		14	4.9	873	24	ABN66184	Streptococcus poly
27		14	4.9	894	14	AAQ39243	FPS DNA. Bacillus
28		14	4.9	894	17	AAT40226	Mutant farnesyldip
29		14	4.9	894	17	AAT40227	Mutant farnesyldip
30		14	4.9	894	17	AAT40229	Native farnesyldip
31		14	4.9	894	19	AAV38455	DNA encoding farne
32		14	4.9	894	19	AAV18516	Bacillus stearothe
33		14	4.9	894	20	AAX86781	Farnesyl diphospha
34		14	4.9	894	21	AAC63809	Bacillus stearothe
35		14	4.9	894	24	ABK96797	B. stearothermophi
36		14	4.9	2155561	24	ABN71527	Streptococcus poly
37		12	4.2	888	23	AAS53505	Haemophilus influe
38		12	4.2	897	20	AAZ33160	Escherichia coli n
39		12	4.2	900	23	AAS52335	E. coli DNA for ce
40		12	4.2	900	23	AAS56349	Salmonella typhi D
41		12	4.2	900	24	AAL40785	DNA relating to th
42		12	4.2	900	24	ABK96781	E.coli prenyl dip
43		12	4.2	903	24	ABQ90268	M. capsulatus gene
44		12	4.2	4390	20	AAZ33182	Escherichia coli D
45	c	12	4.2	11692	22	AAS46249	DNA encoding novel

ALIGNMENTS

RESULT 1
AAS00821
ID AAS00821 standard; DNA; 861 BP.
XX
AC AAS00821;
XX
DT 04-JUL-2001 (first entry)
XX
DE S. aureus HGS072 encoding Farnesyl diphosphatesynthase, IspA.
XX
KW Farnesyl diphosphatesynthase; IspA; immunogen; vaccine; antibody;
KW wound infection; cellulitis; burn infection; eyelid infection;
KW food poisoning; joint infection; neonatal conjunctivitis; osteomyelitis;
KW skin infection; scalded skin syndrome; toxic epidermal necrosis;
KW Ritter's disease; Lyell's disease; toxic shock syndrome; endocarditis;
KW HGS072; ds.
XX
OS Staphylococcus aureus.

```
XX
FH   Key             Location/Qualifiers
FT   CDS             1..861
FT                   /*tag= a
FT                   /product= "IspA"
FT                   /partial
FT                   /note= "No stop codon"
XX
PN   WO200116292-A2.
XX
PD   08-MAR-2001.
XX
PF   31-AUG-2000; 2000WO-US23773.
XX
PR   01-SEP-1999;   99US-0151933.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Choi GH;

DR   WPI; 2001-183259/18.
DR   P-PSDB; AAU00859.
XX
PT   New isolated nucleic acid for use in diagnosing Staphylococcus
PT   infections and in vaccines for eliciting immune responses to the
PT   infections -
XX
PS   Claim 1; Page 23; 225pp; English.
XX
CC   The sequence encodes S. aureus IspA (Farnesyl diphosphatesynthase).
CC   The polynucleotides of the invention are used to detect Staphylococcus
CC   nucleic acids in a biological sample from an animal for diagnosing
CC   Staphylococcus infections. The polypeptides of the invention are used to
CC   detect anti-Staphylococcus antibodies in a biological sample from an
CC   animal to diagnose Staphylococcus infections. The polypeptides are also
CC   used in vaccines to elicit protective antibodies in an animal to a member
CC   of the Staphylococcus genus and for preventing or attenuating an
CC   infection caused by a member of the Staphylococcus genus e.g wound
CC   infection, cellulitis, burn infection, eyelid infection, food poisoning,
CC   joint infection, neonatal conjunctivitis, osteomyelitis, skin infection,
CC   scalded skin syndrome (also known as toxic epidermal necrosis, Ritter's
CC   disease and Lyell's disease), toxic shock syndrome and endocarditis. The
CC   polynucleotides may also be used in vaccines and for preventing or
CC   attenuating a Staphylococcus infection. Antibodies to the polypeptides
CC   may be used to purify, detect and target the polypeptides in in vitro and
CC   in vivo diagnostic and therapeutic methods.
XX
SQ   Sequence 861 BP; 323 A; 114 C; 180 G; 244 T; 0 other;

Alignment Scores:
Pred. No.:              2.79e-280      Length:         861
Score:                  287.00         Matches:        287
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            100.00%        Indels:         0
DB:                     22             Gaps:           0

US-09-925-637-64 (1-287) x AAS00821 (1-861)

Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy         61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy         81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300

Qy        101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600

Qy        201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660

Qy        221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Qy        241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy        261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy        281 LeuLeuGluIleValAspLeu 287
              |||||||||||||||||||||
Db        841 TTATTAGAAATCGTTGATTTA 861

RESULT 2
AAV74466/c
ID   AAV74466 standard; DNA; 1893 BP.
XX
AC   AAV74466;
XX
DT   16-MAR-1999  (first entry)
XX
DE   Staphylococcus aureus contig SEQ ID #155.
XX
KW   Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW   cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW   skin infection; surgical wound infection; scalded skin syndrome;
KW   toxic shock syndrome; ds.
XX
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1261..1320
FT                   /*tag= a
FT                   /note= "these bases represent a line of missing text in
FT                           the sequence listing in the specification. They
FT                           are included to maintain the nucleotide numbering
FT                           given in the specification for this DNA sequence"
XX
PN   EP786519-A2.
XX
```

PD 30-JUL-1997.
XX
PF 07-JAN-1997; 97EP-0100117.
XX
PR 05-JAN-1996; 96US-0009861.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
PI Rosen CA;
XX
DR WPI; 1997-374922/35.
XX
PT Polynucleotide(s) and proteins derived from Staphylococcus aureus -
PT stored on computer readable medium and used in the production of
PT anti-S.aureus vaccines
XX
PS Claim 1; Page 784-785; 3271pp; English.
XX
CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the S.aureus DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against S.aureus infection. The
CC polypeptides can also be used in a kit for the immunodetection of
CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC skin and surgical wound infections, scalded skin syndrome, toxic shock
CC syndrome, etc. Organisms transformed with the DNA sequences can be used
CC for recombinant production of the polypeptides. The new DNA sequences
CC (and their fragments) are useful as primers or probes for isolating
CC homologues of any of the S.aureus DNA sequences contained on the
CC computer readable medium.
XX
SQ Sequence 1893 BP; 552 A; 336 C; 233 G; 708 T; 64 other;

Alignment Scores:

Pred. No.:	9.02e-262	Length:	1893
Score:	269.00	Matches:	269
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	93.73%	Indels:	0
DB:	18	Gaps:	0

US-09-925-637-64 (1-287) x AAV74466 (1-1893)

```
Qy         19 ValAlaIleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeu 38
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GTTGCGATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTA 1201

Qy         39 AsnAlaGlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsn 58
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 AATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAAT 1141

Qy         59 ThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSer 78
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 ACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCA 1081

Qy         79 LeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThr 98
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACA 1021

Qy         99 AsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLys 118
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 AATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAA 961

Qy        119 AlaPheGluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeu 138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 GCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTA 901

Qy        139 GlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMet 158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 CAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATG 841

Qy        159 GlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThr 178
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 CAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACA 781

Qy        179 GlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThr 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GGAGCATTATTAACTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACA 721

Qy        199 ThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAsp 218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACTAAAGAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGAT 661

Qy        219 AspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeu 238
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 GATTTATTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTT 601

Qy        239 GluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeu 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GAAAATAATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTG 541

Qy        259 ThrTyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThr 278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ACTTATCATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACA 481

Qy        279 LysHisLeuLeuGluIleValAspLeu 287
              |||||||||||||||||||||||||||
Db        480 AAACACTTATTAGAAATCGTTGATTTA 454
```

RESULT 3
AAS54848
ID AAS54848 standard; DNA; 882 BP.
XX
AC AAS54848;
XX
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus DNA for cellular proliferation protein #1160.
XX
KW Antisense; ds; prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR P-PSDB; AAU36989.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX

```
PS   Claim 27; Seq ID No 8485; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence encodes an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 882 BP; 330 A; 116 C; 185 G; 251 T; 0 other;

Alignment Scores:
Pred. No.:                2.32e-189      Length:          882
Score:                    197.00         Matches:         268
Percent Similarity:       99.26%         Conservative:    0
Best Local Similarity:    99.26%         Mismatches:      1
Query Match:              68.64%         Indels:          2
DB:                       23             Gaps:            0

US-09-925-637-64 (1-287) x AAS54848 (1-882)

Qy        1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy       21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120

Qy       41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy       61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy       81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCAT 300

Qy      101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy      121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy      141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy      161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy      181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThr-ThrLy 200
            ||||||||||||||||||||||||||||||||||||||||||||||||||    |||||
Db      541 CTATTAACATTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGC-AACTAA 599

Qy      200 sGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLe 220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      600 AGAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTTCAGATTAAAGATGATTT 659

Qy      220 uLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAs 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      660 ATTAGACTGCTATGGTGATGAAGCGAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAA 719

Qy      240 nAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTy 260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      720 TAATAAAAGTACATACGTGAGTTTATTAGGAAAAGATGGCGCAGAAGATAAATTGACTTA 779

Qy      260 rHisArgAspAlaAlaValAspGluLeu 269
            |||||||||||||||||||||||||||
Db      780 TCATAGAGACGCAGCAGTGGATGAACTA 807

RESULT 4
AAS51602
ID   AAS51602 standard; DNA; 864 BP.
XX
AC   AAS51602;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus DNA for cellular proliferation protein #19.
XX
KW   Antisense; ds; prokaryotic cellular proliferation gene;
KW   antibiotic; antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   P-PSDB; AAU33743.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Claim 27; Seq ID No 4184; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence encodes an
```

```
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 864 BP; 323 A; 112 C; 182 G; 247 T; 0 other;

Alignment Scores:
Pred. No.:              2.56e-184       Length:         864
Score:                  192.00          Matches:        263
Percent Similarity:     99.25%          Conservative:   0
Best Local Similarity:  99.25%          Mismatches:     1
Query Match:            66.90%          Indels:         2
DB:                     23              Gaps:           0

US-09-925-637-64 (1-287) x AAS51602 (1-864)

Qy          6 MetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerVal 25
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTA 60

Qy         26 MetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIle 45
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATC 120

Qy         46 ArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMet 65
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATG 180

Qy         66 LysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCA 240

Qy         86 AlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGlu 105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCATAAAGTATATGGTGAG 300

Qy        106 TrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSer 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGT 360

Qy        126 AspAspArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSer 145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGT 420

Qy        146 GlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIle 165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATT 480

Qy        166 AspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAla 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCACTATTAACATTTGCG 540

Qy        186 ValMetSerAlaAlaAspIleAlaAsnValAspAspThr-ThrLysGluHisLeuGluSe 205
              ||||||||||||||||||||||||||||||||||||    ||||||||||||||||||||
Db        541 GTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGC-AACTAAAGAACATTTAGAAAG 599

Qy        205 rTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGl 225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 TTATAGTTATCATTTAGGTATGATGTTTCAGATTAAAGATGATTTATTAGACTGCTATGG 659

Qy        225 yAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTy 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 TGATGAAGCGAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTACATA 719

Qy        245 rValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAl 265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 CGTGAGTTTATTAGGAAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGACGCAGC 779

Qy        265 aValAspGluLeu 269
              |||||||||||||
Db        780 AGTGGATGAACTA 792
```

```
RESULT 5
AAS50245/c
ID   AAS50245 standard; DNA; 413 BP.
XX
AC   AAS50245;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus cellular proliferation inhibitory sequence #1469.
XX
KW   Antisense; ss; prokaryotic cellular proliferation;
KW   antibiotic; antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Claim 1; Seq ID No 2822; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence is an antisense
CC   oligonucleotide of the invention.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 413 BP; 115 A; 89 C; 58 G; 151 T; 0 other;

Alignment Scores:
Pred. No.:              4.55e-129       Length:         413
Score:                  137.00          Matches:        137
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            47.74%          Indels:         0
DB:                     23              Gaps:           0

US-09-925-637-64 (1-287) x AAS50245 (1-413)
```

```
Qy          54 LeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMet 73
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         413 TTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATG 354

Qy          74 IleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArg 93
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         353 ATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGA 294

Qy          94 ArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAsp 113
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         293 CGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGAT 234

Qy         114 AlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAspArgLeuThrAspGluVal 133
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         233 GCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTA 174

Qy         134 LysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGly 153
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         173 AAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGT 114

Qy         154 GlnMetLeuAspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIle 173
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         113 CAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATA 54

Qy         174 HisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAla 190
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          53 CACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCA 3
```

RESULT 6
AAS50698/c
ID AAS50698 standard; DNA; 337 BP.
XX
AC AAS50698;
XX
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation inhibitory sequence #1922.
XX
KW Antisense; ss; prokaryotic cellular proliferation;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.

21-MAR-2001; 2001WO-US09180.

PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Claim 1; Seq ID No 3275; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence is an antisense
CC oligonucleotide of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 337 BP; 96 A; 71 C; 47 G; 123 T; 0 other;

Alignment Scores:

Pred. No.:	6.7e-104	Length:	337
Score:	112.00	Matches:	112
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	39.02%	Indels:	0
DB:	23	Gaps:	0

US-09-925-637-64 (1-287) x AAS50698 (1-337)

```
Qy          76 ThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGly 95
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         336 ACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGA 277

Qy          96 LysLeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeu 115
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         276 AAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTA 217

Qy         116 LeuThrLysAlaPheGluLeuIleSerSerAspAspArgLeuThrAspGluValLysIle 135
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         216 TTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATA 157

Qy         136 LysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMet 155
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         156 AAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATG 97

Qy         156 LeuAspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLys 175
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          96 TTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAA 37

Qy         176 ThrLysThrGlyAlaLeuLeuThrPheAlaValMet 187
               ||||||||||||||||||||||||||||||||||||
Db          36 ACAAAAACAGGAGCATTATTAACTTTTGCGGTTATG 1
```

RESULT 7
AAA92031
ID AAA92031 standard; DNA; 882 BP.
XX
AC AAA92031;
XX
DT 12-JAN-2001 (first entry)
XX
DE Staphylococcus aureus ispA coding sequence.
XX
KW IspA; bacterial disease; respiratory tract infection;
KW gastrointestinal infection; cardiac infection; Helicobacter pylori;
KW stomach cancer; stomach ulcer; gastritis; ds.
XX
OS Staphylococcus aureus.
XX

FH	Key	Location/Qualifiers
FT	CDS	1..882
FT		/*tag= a
FT		/product= "IspA"

XX
PN US6107058-A.
XX
PD 22-AUG-2000.
XX
PF 26-MAR-1999; 99US-0276873.
XX
PR 26-MAR-1999; 99US-0276873.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Gwynn M, Wilding EI;
XX
DR WPI; 2000-578535/54.
DR P-PSDB; AAB23333.
XX
PT Novel farnesyl diphosphate synthase polynucleotide from staphylococcus
PT aureus useful for diagnosis and treatment of bacterial infections and
PT as hybridization probe for isolating genomic clones -
XX
[redacted] Claim 1; column 1-4; 15pp; English.

CC The present sequence is the coding sequence for the Staphylococcus aureus
CC IspA protein. This gene and the protein it encodes can be used in many
CC research assays, as well as treatments for bacterial diseases such as
CC infections of the respiratory tract (including otitis media, bacterial
CC tracheitis, acute epiglottitis, thyroiditis, empyema and lung abscesses),
CC cardiac infections such as infective endocarditis, gastrointestinal
CC infections including secretory diarrhoea, splenic abscesses and
CC retroperitoneal abscesses, CNS infections such as cerebral abscesses, eye
CC infections (including blepharitis, conjunctivitis, keratitis,
CC endophthalmitis, preseptal and orbital cellulitis and darcryocystitis),
CC kidney and urinary tract infections such as epididymitis, intrarenal and
CC perinephric abscesses and toxic shock syndrome, skin diseases (including
CC impetigo, folliculitis, cutaneous abscesses, wound infection and
CC bacterial myositis), bone and joint infections such as septic
CC arthritis and osteomyelitis, septic thrombophlebitis, food poisoning and
CC scalded skin syndrome. In addition, they can be used to treat
CC diseases caused by Helicobacter pylori, including stomach cancer, stomach
CC ulcers and gastritis.
XX
SQ Sequence 882 BP; 328 A; 118 C; 186 G; 250 T; 0 other;

Alignment Scores:

Pred. No.:	2.13e-93	Length:	882
Score:	102.00	Matches:	284
Percent Similarity:	97.93%	Conservative:	0
Best Local Similarity:	97.93%	Mismatches:	3
Query Match:	35.54%	Indels:	6
[redacted]:	21	Gaps:	0

US-09-925-637-64 (1-287) x AAA92031 (1-882)

```
Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGCATGTTGTATTCATTAAATGCT 120

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy         61 TyrGluLeuGlyMet-LysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIl 80
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTGT-GAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTAT 239

Qy         80 eHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGly-LysLeuThrAsnH 100
              ||||||||||||||||||||||||||||||||||||||||||||    ||||||||||||||
Db        240 TCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAAG-GAAATTAACAAATC 298

Qy        100 isLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCAT 358

Qy        120 heGluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnA 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 TTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAAC 418

Qy        140 rgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnS 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 GGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAA 478

Qy        160 erGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        479 GCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAG 538

Qy        180 laLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThr-Thr 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||    ||| |||
Db        539 CATTATTAACTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGC-AACT 597

Qy        200 LysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAsp 219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 AAAGAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGAT 657

Qy        220 LeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGlu 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 TTATTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAA 717

Qy        240 AsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThr 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 AATAATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACT 777

Qy        260 TyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLys 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 TATCATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAA 837

Qy        280 HisLeuLeuGluIleValAspLeu 287
              |||||||||||||||||||||||||
Db        838 CACTTATTAGAAATCGTTGATTTA 861
```

RESULT 8
AAH52415
ID AAH52415 standard; DNA; 783 BP.
XX
AC AAH52415;
XX
DT 03-SEP-2001 (first entry)
XX
DE S. epidermidis open reading frame nucleotide sequence SEQ ID NO:223.
XX
KW Staphylococcus epidermidis SR1 strain; infection; diagnosis;
KW vaccination; endocarditis; ds.
XX
OS Staphylococcus epidermidis.
XX
PN WO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US30782.
XX
PR 09-NOV-1999; 99US-0164258.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Kimmerly WJ;
XX
DR WPI; 2001-316495/33.
DR P-PSDB; AAG81565.
XX
PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis -
XX

PS Claim 8; Page 103; 2188pp; English.
XX
CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ Sequence 783 BP; 297 A; 93 C; 147 G; 246 T; 0 other;

Alignment Scores:

Pred. No.:	8.64e-28	Length:	783
Score:	37.00	Matches:	37
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.89%	Indels:	0
DB:	22	Gaps:	0

US-09-925-637-64 (1-287) x AAH52415 (1-783)

```
Qy         70 AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsn 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        109 GCATTGGAAATGATTCATACTTATTCTTTAATTCATGATGATTTACCAGCAATGGATAAT 168

Qy         90 AspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrp 106
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 GACGATTACCGTAGAGGAAAATTAACAAATCATAAAGTTTATGGTGAATGG 219
```

RESULT 9
ABN90871
ID ABN90871 standard; DNA; 909 BP.
XX
AC ABN90871;

DT 24-JUL-2002 (first entry)
XX
DE Staphylococcus epidermidis ORF nucleic acid sequence SEQ ID NO:334.
XX
KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW antibacterial; gene therapy; gene; ds.
XX
OS Staphylococcus epidermidis.
XX
PN US6380370-B1.
XX
PD 30-APR-2002.
XX
PF 13-AUG-1998; 98US-0134001.
XX
PR 14-AUG-1997; 97US-055779P.
PR 08-NOV-1997; 97US-064964P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Doucette-Stamm LA, Bush D;
XX
DR WPI; 2002-381255/41.
DR P-PSDB; ABP38326.
XX
PT Novel isolated nucleic acid encoding a Staphylococcus epidermidis
PT polypeptide, useful for diagnosing and treating bacterial infections -
XX
PS Disclosure; SEQ ID 334; 267pp; English.
XX
CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC antibacterial activity and can be used in gene therapy. The sequences
CC can also be used in the diagnosis and treatment of bacterial infections,
CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life
CC cycle or inhibit S. epidermidis infection.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site.
XX
SQ Sequence 909 BP; 361 A; 106 C; 165 G; 277 T; 0 other;

Alignment Scores:

Pred. No.:	9.97e-28	Length:	909
Score:	37.00	Matches:	37
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.89%	Indels:	0
DB:	24	Gaps:	0

US-09-925-637-64 (1-287) x ABN90871 (1-909)

```
Qy         70 AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsn 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 GCATTGGAAATGATTCATACTTATTCTTTAATTCATGATGATTTACCAGCAATGGATAAT 294

Qy         90 AspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrp 106
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 GACGATTACCGTAGAGGAAAATTAACAAATCATAAAGTTTATGGTGAATGG 345
```

RESULT 10
AAH54780
ID AAH54780 standard; DNA; 3444 BP.
XX
AC AAH54780;
XX
DT 03-SEP-2001 (first entry)
XX
DE S. epidermidis genomic polynucleotide sequence SEQ ID NO:4144.
XX
KW Staphylococcus epidermidis SR1 strain; infection; diagnosis;
KW vaccination; endocarditis; ds.
XX
OS Staphylococcus epidermidis.
XX
PN WO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US30782.
XX
PR 09-NOV-1999; 99US-0164258.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Kimmerly WJ;
XX
DR WPI; 2001-316495/33.
XX
PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis -
XX
PS Claim 8; Page 1844-1845; 2188pp; English.
XX
CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides

CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.
XX

Sequence 3444 BP; 1365 A; 451 C; 571 G; 1057 T; 0 other;

Alignment Scores:

Pred. No.:	3.58e-27	Length:	3444
Score:	37.00	Matches:	37
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.89%	Indels:	0
DB:	22	Gaps:	0

US-09-925-637-64 (1-287) x AAH54780 (1-3444)

```
Qy         70 AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsn 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        973 GCATTGGAAATGATTCATACTTATTCTTTAATTCATGATGATTTACCAGCAATGGATAAT 1032

Qy         90 AspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrp 106
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1033 GACGATTACCGTAGAGGAAAATTAACAAATCATAAAGTTTATGGTGAATGG 1083
```

RESULT 11
AAH54634/c
ID AAH54634 standard; DNA; 3454 BP.
XX
AC AAH54634;
XX
DT 03-SEP-2001 (first entry)
XX
S. epidermidis genomic polynucleotide sequence SEQ ID NO:3998.

Staphylococcus epidermidis SR1 strain; infection; diagnosis;
KW vaccination; endocarditis; ds.
XX
OS Staphylococcus epidermidis.
XX
PN WO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US30782.
XX
PR 09-NOV-1999; 99US-0164258.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Kimmerly WJ;
XX
DR WPI; 2001-316495/33.
XX
PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis -
XX
PS Claim 8; Page 1671-1672; 2188pp; English.
XX
CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ Sequence 3454 BP; 1005 A; 574 C; 493 G; 1382 T; 0 other;

Alignment Scores:

Pred. No.:	3.59e-27	Length:	3454
Score:	37.00	Matches:	37
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.89%	Indels:	0
DB:	22	Gaps:	0

US-09-925-637-64 (1-287) x AAH54634 (1-3454)

```
Qy         70 AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsn 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 GCATTGGAAATGATTCATACTTATTCTTTAATTCATGATGATTTACCAGCAATGGATAAT 91

Qy         90 AspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrp 106
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 GACGATTACCGTAGAGGAAAATTAACAAATCATAAAGTTTATGGTGAATGG 40
```

RESULT 12
AAH54565/c
ID AAH54565 standard; DNA; 4045 BP.
XX
AC AAH54565;
XX
DT 03-SEP-2001 (first entry)
XX
DE S. epidermidis genomic polynucleotide sequence SEQ ID NO:3929.
XX
KW Staphylococcus epidermidis SR1 strain; infection; diagnosis;
KW vaccination; endocarditis; ds.
XX
OS Staphylococcus epidermidis.
XX
PN WO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US30782.
XX
PR 09-NOV-1999; 99US-0164258.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Kimmerly WJ;
XX
DR WPI; 2001-316495/33.
XX
PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis -

```
XX
PS   Claim 8; Page 1586-1588; 2188pp; English.
XX
CC   AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC   (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC   (I) and (II) can have antibacterial activity and therefore can be used
CC   in vaccination. The nucleic acids (I) may be used to produce the
CC   S. epidermidis polypeptides (II) via the production of vectors
CC   containing them which are used to produce hosts cells which express the
CC   polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC   used to vaccinate subjects and to raise antibodies against the bacteria.
CC   The polypeptides may also be used to assay for other inhibitors of their
CC   activity and therefore identify compounds that may be used for the
CC   treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC   AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC   polynucleotide sequences from the present invention. AAH55091 to
CC   AAH55098 represent oligonucleotide sequences and primers which are used
CC   in the exemplification of the present invention.
CC   N.B. The present invention specifically claims all the polynucleotide
CC   sequences given in the sequence listing of the present specification,
CC   however the sequence listing only goes up to SEQ ID NO:4454 so even
CC   though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC   no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ   Sequence 4045 BP; 1249 A; 691 C; 536 G; 1569 T; 0 other;

Alignment Scores:
Pred. No.:                4.18e-27        Length:         4045
Score:                    37.00           Matches:        37
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              12.89%          Indels:         0
DB:                       22              Gaps:           0

US-09-925-637-64 (1-287) x AAH54565 (1-4045)

Qy          70 AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsn 89
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2106 GCATTGGAAATGATTCATACTTATTCTTTAATTCATGATGATTTACCAGCAATGGATAAT 2047

Qy          90 AspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrp 106
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2046 GACGATTACCGTAGAGGAAAATTAACAAATCATAAAGTTTATGGTGAATGG 1996

RESULT 13
AAV15824
ID   AAV15824 standard; DNA; 486 BP.

     AAV15824;

DT   05-JUN-1998  (first entry)
XX
DE   Probe B500 for Micrococcus luteus HexPS gene.
XX
KW   HexPS gene; hex1; hex2; hex3; prenyl diphosphate synthetase;
KW   subunit A; subunit B; polyprenyl diphosphate; vitamin K;
KW   ubiquinone; probe; ss.
XX
OS   Micrococcus luteus.
XX
PN   EP812914-A2.
XX
PD   17-DEC-1997.
XX
PF   13-JUN-1997;   97EP-0109692.
XX
PR   14-JUN-1996;   96JP-0154441.
XX
PA   (TOYT ) TOYOTA JIDOSHA KK.
XX
PI   Cho Y,  Koike A,  Koyama T,  Muramatsu M,  Ogura K;
PI   Shimizu N;
XX
DR   WPI; 1998-034975/04.
XX
PT   DNA encoding prenyl diphosphate synthetase subunit(s) - new
PT   Micrococcus prenyl diphosphate synthetase subunit polypeptide(s),
PT   and methods for preparing enzymes from subunit(s)
XX
PS   Disclosure; Pages 30-31; 46pp; English.
XX
CC   The present sequence is a probe for the Micrococcus luteus HexPS
CC   gene, which comprises three open reading frames designated hex1,
CC   hex2 and hex3. hex1 and hex3 are claimed in claims 3 and 3, and
CC   encode prenyl diphosphate synthetase (PDS) subunit A and B,
CC   respectively.
CC   Substances synthesised by PDS, i.e. polyprenyl diphosphates, are
CC   precursors of physiologically active substances, e.g. vitamin K and
CC   ubiquinones.
XX
SQ   Sequence 486 BP; 148 A; 85 C; 98 G; 149 T; 6 other;

Alignment Scores:
Pred. No.:                6e-24           Length:         486
Score:                    33.00           Matches:        33
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              11.50%          Indels:         0
DB:                       19              Gaps:           0

US-09-925-637-64 (1-287) x AAV15824 (1-486)

Qy          69 IleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAsp 88
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          85 ATTGCGCTTGAAATGATTCATACGTATTCTTTAATTCACGATGACTTGCCGGCAATGGAT 144

Qy          89 AsnAspAspTyrArgArgGlyLysLeuThrAsnHisLys 101
               |||||||||||||||||||||||||||||||||||||||
Db         145 AATGATGACTATCGTCGCGGTAAACTCACGAATCATAAG 183

RESULT 14
ABQ71010
ID   ABQ71010 standard; DNA; 5635 BP.
XX
AC   ABQ71010;
XX
DT   29-AUG-2002  (first entry)
XX
DE   Listeria monocytogenes 4b contig DNA sequence #952.
XX
KW   Antibacterial; Listeria; food contamination; mutational analysis;
KW   infection; ds.
XX
OS   Listeria monocytogenes 4b.
XX
PN   WO200228891-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-FR03061.
XX
PR   04-OCT-2000; 2000FR-0012697.
XX
PA   (INSP ) INST PASTEUR.
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Kunst F,  Glaser P;
XX
DR   WPI; 2002-332479/37.
XX
PT   New genomic sequences from Listeria species, useful for detection,
PT   treatment and prevention of infection, also related polypeptides,
PT   antibodies and modulators  -
XX
PS   Claim 14; SEQ ID 3823; 180pp; French.
XX
```

```
CC   The present invention relates to nucleic acid sequences
CC   (ABQ67188-ABQ71212) from Listeria sp. The sequences are useful as probes
CC   and primers for identification and/or detection of Listeria (e.g. as
CC   contaminants in foods, or mutational analysis) and for analysis of
CC   gene expression. Proteins encoded by the nucleic acid sequences can be
CC   used to screen for compounds that modulate gene expression, replication
CC   and pathogenicity of Listeria (potential therapeutic agents), also for
CC   treating infections by Listeria, and are useful as immunogens in
CC   anti-Listeria vaccines.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 5635 BP; 1896 A; 927 C; 1238 G; 1574 T; 0 other;

Alignment Scores:
Pred. No.:              7.25e-17      Length:          5635
Score:                  27.00         Matches:         27
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
  ery Match:            9.41%         Indels:          0
  :                     24            Gaps:            0

US-09-925-637-64 (1-287) x ABQ71010 (1-5635)

Qy         70 AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsn 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3091 GCGTTAGAAATGATTCATACGTATAGCCTGATCCATGATGATTTACCAGCAATGGATAAT 3150

Qy         90 AspAspTyrArgArgGlyLys 96
              ||||||||||||||||||||
Db       3151 GATGACTATCGTCGCGGCAAA 3171

RESULT 15
ABA03041
ID   ABA03041 standard; DNA; 2944528 BP.
XX
AC   ABA03041;
XX
DT   05-FEB-2002  (first entry)
XX
DE   Listeria monocytogenes EGD-e genome sequence.
XX
KW   Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW   vitamin B12; bacterial infection; disease; ds.
XX
OS   Listeria monocytogenes.
XX
PN   WO200177335-A2.
XX
PD   18-OCT-2001.
XX
PF   11-APR-2001; 2001WO-FR01118.
XX
PR   11-APR-2000; 2000FR-0004629.
XX
PA   (INSP ) INST PASTEUR.
XX
PI   Buchrieser C,  Frangeul L,  Couve E,  Rusniok C,  Fsihi H,  Dehoux P;
PI   Dussurget O,  Chetouani F,  Nedjari H,  Glaser P,  Kunst F,  Cossart P;
PI   Daniels J,  Goebel W,  Kreft J,  Kuhn M,  Ng E,  Vazquez-Boland JA;
PI   Dominguez-Bernal G,  Garrido-Garcia P,  Tierrez-Martinez A,  Amend A;
PI   Chakraborty T,  Domann E,  Hain T,  Berche P,  Charbit A,  Durant L;
PI   Perez-Diaz J,  Baquero F,  Garcia Del Portillo F,  Gomez-Lopez N;
PI   Maduenio E,  De Pablos B,  Wehland J,  Kaerst U,  Entian K,  Hauf J;
PI   Rose M,  Voss H;
XX
DR   WPI; 2002-010914/01.
XX
PT   Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT   and prevention of Listeria and related bacterial infections, and
PT   related polypeptides  -
XX
PS   Claim 1; SEQ ID No 1; 192pp; French.
XX
CC   The present sequence is the genome sequence of Listeria monocytogenes
CC   EGD-e. This sequence and fragments of this sequence are useful for
CC   selecting probes and primers for detecting genes in L. monocytogenes and
CC   related organisms, and to study genetic polymorphisms and other genomes.
CC   Proteins (ABB47297-ABB50149) expressed from the present sequence are
CC   useful for raising specific antibodies, identification of L.
CC   monocytogenes and related organisms, and for biosynthesis and
CC   biodegradation, especially biosynthesis of Vitamin B12. This sequence and
CC   proteins encoded by it are also useful for selecting compounds that
CC   regulate gene expression and cell replication and modulate L.
CC   monocytogenes-related diseases. In addition, this sequence and proteins
CC   encoded by it are useful in pharmaceutical and vaccines compositions for
CC   the treatment or prevention of infections by L. monocytogenes and related
CC   organisms.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2944528 BP; 914202 A; 563301 C; 555061 G; 911964 T; 0 other;

Alignment Scores:
Pred. No.:              2.94e-14      Length:          2944528
Score:                  27.00         Matches:         27
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            9.41%         Indels:          0
DB:                     24            Gaps:            0

US-09-925-637-64 (1-287) x ABA03041 (1-2944528)

Qy         70 AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsn 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1386147 GCGTTAGAAATGATTCACACATACAGCCTTATTCATGATGATCTACCAGCAATGGATAAC 1386206

Qy         90 AspAspTyrArgArgGlyLys 96
              ||||||||||||||||||||
Db    1386207 GATGATTATCGTCGAGGCAAG 1386227


Search completed: May 30, 2003, 22:59:05
Job time : 722 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 30, 2003, 22:36:16 ; Search time 59 Seconds
(without alignments)
1491.801 Million cell updates/sec

Title: US-09-925-637-64
Perfect score: 287
Sequence: 1 MTNLPMNKLIDEVNNELSVA.....ELTQIDEQFNTKHLLEIVDL 287

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

arched: 441362 seqs, 153338381 residues

Word size: 1

Total number of hits satisfying chosen parameters: 878600

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09925637/runat_23052003_174850_24168/app_query.fasta_1.455
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=oli.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09925637_@CGN_1_1_40_@runat_23052003_174850_24168 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	102	35.5	882	3	US-09-276-873-1	Sequence 1, Appli
2	37	12.9	909	4	US-09-134-001C-334	Sequence 334, App
3	33	11.5	486	3	US-09-217-609A-28	Sequence 28, Appl
4	33	11.5	486	4	US-08-873-235B-28	Sequence 28, Appl
5	14	4.9	893	1	US-08-333-321-1	Sequence 1, Appli
6	14	4.9	894	1	US-08-534-910B-2	Sequence 2, Appli
7	14	4.9	894	1	US-08-534-910B-3	Sequence 3, Appli
8	14	4.9	894	1	US-08-534-910B-5	Sequence 5, Appli
9	14	4.9	894	3	US-08-886-466-1	Sequence 1, Appli
10	14	4.9	894	4	US-09-475-304-1	Sequence 1, Appli
11	14	4.9	894	4	US-09-101-126-4	Sequence 4, Appli
12	14	4.9	894	4	US-09-367-528A-4	Sequence 4, Appli
13	11	3.8	876	3	US-09-275-742-1	Sequence 1, Appli
14	11	3.8	1284	4	US-08-858-207A-173	Sequence 173, App
c 15	11	3.8	10011	4	US-08-961-527-76	Sequence 76, Appl
16	9	3.1	33	2	US-08-898-560-13	Sequence 13, Appl
17	9	3.1	972	1	US-08-506-404D-3	Sequence 3, Appli
18	9	3.1	972	3	US-09-035-754-3	Sequence 3, Appli
19	8	2.8	36	2	US-08-898-560-7	Sequence 7, Appli
20	8	2.8	861	3	US-09-187-050-4	Sequence 4, Appli
21	8	2.8	885	3	US-09-187-050-11	Sequence 11, Appl
22	8	2.8	894	1	US-08-534-910B-1	Sequence 1, Appli
23	8	2.8	894	1	US-08-534-910B-4	Sequence 4, Appli
24	8	2.8	894	4	US-09-367-528A-2	Sequence 2, Appli
25	8	2.8	1179	3	US-09-187-050-13	Sequence 13, Appl
26	8	2.8	1179	3	US-09-187-050-15	Sequence 15, Appl
27	8	2.8	1179	3	US-09-187-050-17	Sequence 17, Appl
28	8	2.8	1179	3	US-09-187-050-19	Sequence 19, Appl
29	8	2.8	1179	3	US-09-187-050-21	Sequence 21, Appl
30	8	2.8	1179	3	US-09-187-050-23	Sequence 23, Appl
31	8	2.8	1179	3	US-09-187-050-25	Sequence 25, Appl
32	8	2.8	1426	2	US-08-284-465-2	Sequence 2, Appli
33	8	2.8	1426	2	US-08-284-465-7	Sequence 7, Appli
34	8	2.8	1889	3	US-09-187-050-1	Sequence 1, Appli
35	8	2.8	4061	3	US-08-425-843-1	Sequence 1, Appli
36	7	2.4	34	3	US-08-886-466-3	Sequence 3, Appli
37	7	2.4	34	3	US-08-886-466-4	Sequence 4, Appli
38	7	2.4	34	3	US-08-886-466-5	Sequence 5, Appli
39	7	2.4	34	3	US-08-886-466-6	Sequence 6, Appli
40	7	2.4	34	3	US-08-886-466-7	Sequence 7, Appli
41	7	2.4	34	3	US-08-886-466-8	Sequence 8, Appli
42	7	2.4	34	3	US-08-886-466-9	Sequence 9, Appli
43	7	2.4	34	4	US-09-475-304-3	Sequence 3, Appli
44	7	2.4	34	4	US-09-475-304-4	Sequence 4, Appli
45	7	2.4	34	4	US-09-475-304-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-276-873-1
; Sequence 1, Application US/09276873
; Patent No. 6107058
; GENERAL INFORMATION:
; APPLICANT: Wilding, Edwina Imogen
; APPLICANT: Gwynn, Michael
; TITLE OF INVENTION: IspA
; FILE REFERENCE: GM10208
; CURRENT APPLICATION NUMBER: US/09/276,873
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 882
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-276-873-1

Alignment Scores:

Pred. No.:	1.87e-99	Length:	882
Score:	102.00	Matches:	284
Percent Similarity:	97.93%	Conservative:	0
Best Local Similarity:	97.93%	Mismatches:	3
Query Match:	35.54%	Indels:	6
DB:	3	Gaps:	0

US-09-925-637-64 (1-287) x US-09-276-873-1 (1-882)

```
Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGCATGTTGTATTCATTAAATGCT 120
```

```
Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy         61 TyrGluLeuGlyMet-LysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIl 80
              |||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTGT-GAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTAT 239

Qy         80 eHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGly-LysLeuThrAsnH 100
              |||||||||||||||||||||||||||||||||||||||||||    |||||||||||||
Db        240 TCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAAG-GAAATTAACAAATC 298

Qy        100 isLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCAT 358

Qy        120 heGluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnA 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 TTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAAC 418

Qy        140 rgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnS 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 GGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAA 478

Qy        160 erGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        479 GCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAG 538

Qy        180 laLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThr-Thr 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||    ||| 
Db        539 CATTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGC-AACT 597

Qy        200 LysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAsp 219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 AAAGAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGAT 657

Qy        220 LeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGlu 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 TTATTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAA 717

Qy        240 AsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThr 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 AATAATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACT 777

Qy        260 TyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLys 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 TATCATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAA 837

Qy        280 HisLeuLeuGluIleValAspLeu 287
              |||||||||||||||||||||||
Db        838 CACTTATTAGAAATCGTTGATTTA 861

RESULT 2
US-09-134-001C-334
; Sequence 334, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 334
;   LENGTH: 909
;   TYPE: DNA
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-334

Alignment Scores:
Pred. No.:                3.04e-30      Length:          909
Score:                    37.00         Matches:         37
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              12.89%        Indels:          0
DB:                       4             Gaps:            0

US-09-925-637-64 (1-287) x US-09-134-001C-334 (1-909)

Qy         70 AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsn 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 GCATTGGAAATGATTCATACTTATTCTTTAATTCATGATGATTTACCAGCAATGGATAAT 294

Qy         90 AspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrp 106
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 GACGATTACCGTAGAGGAAAATTAACAAATCATAAAGTTTATGGTGAATGG 345

RESULT 3
US-09-217-609A-28
; Sequence 28, Application US/09217609A
; Patent No. 6071733
;  GENERAL INFORMATION:
;    APPLICANT:  MURAMATSU, Masayoshi
;    APPLICANT:  KOIKE, Ayumi
;    APPLICANT:  OGURA, Kyozo
;    APPLICANT:  KOYAMA, Tanetoshi
;    APPLICANT:  SHIMIZU, Naoto
;    APPLICANT:  CHO, Yenwin
;    TITLE OF INVENTION:  Prenyl Diphosphate Synthetase Genes
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenyon & Kenyon
;      STREET:  1025 Connecticut Avenue, NW - Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  US
;      ZIP:  20036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;      SOFTWARE:  WordPerfect 6.1 for Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/217,609A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/873,235
;      FILING DATE:  11-Jun-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  TOFFENETTI, Judith L.
;      REGISTRATION NUMBER:  39,048
;      REFERENCE/DOCKET NUMBER:  10235/1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-429-1776
;      TELEFAX:  202-429-0796
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  486 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  genomic DNA
US-09-217-609A-28

Alignment Scores:
Pred. No.:                2.96e-26      Length:          486
Score:                    33.00         Matches:         33
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
```

```
Query Match:           11.50%      Indels:        0
DB:                    3           Gaps:          0

US-09-925-637-64 (1-287) x US-09-217-609A-28 (1-486)

Qy         69 IleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAsp 88
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 ATTGCGCTTGAAATGATTCATACGTATTCTTTAATTCACGATGACTTGCCGGCAATGGAT 144

Qy         89 AsnAspAspTyrArgArgGlyLysLeuThrAsnHisLys 101
              |||||||||||||||||||||||||||||||||||||
Db        145 AATGATGACTATCGTCGCGGTAAACTCACGAATCATAAG 183

RESULT 4
US-08-873-235B-28
; Sequence 28, Application US/08873235B
; Patent No. 6174715
;  GENERAL INFORMATION:
;    APPLICANT:  MURAMATSU, Masayoshi
     APPLICANT:  KOIKE, Ayumi
     APPLICANT:  OGURA, Kyozo
     APPLICANT:  KOYAMA, Tanetoshi
,    APPLICANT:  SHIMIZU, Naoto
;    APPLICANT:  CHO, Yenwin
;    TITLE OF INVENTION:  Prenyl Diphosphate Synthetase Genes
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenyon & Kenyon
;      STREET:  1025 Connecticut Avenue, NW - Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  US
;      ZIP:  20036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;      SOFTWARE:  WordPerfect 6.1 for Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/873,235B
;      FILING DATE:  11-Jun-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 154441/1996
;      FILING DATE:  14-Jun-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  TOFFENETTI, Judith L.
;      REGISTRATION NUMBER:  39,048
       REFERENCE/DOCKET NUMBER:  10235/1
     TELECOMMUNICATION INFORMATION:
       TELEPHONE:  202-429-1776
;      TELEFAX:  202-429-0796
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  486 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  genomic DNA
US-08-873-235B-28

Alignment Scores:
Pred. No.:             2.96e-26    Length:        486
Score:                 33.00       Matches:       33
Percent Similarity:    100.00%     Conservative:  0
Best Local Similarity: 100.00%     Mismatches:    0
Query Match:           11.50%      Indels:        0
DB:                    4           Gaps:          0

US-09-925-637-64 (1-287) x US-08-873-235B-28 (1-486)

Qy         69 IleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAsp 88
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 ATTGCGCTTGAAATGATTCATACGTATTCTTTAATTCACGATGACTTGCCGGCAATGGAT 144

Qy         89 AsnAspAspTyrArgArgGlyLysLeuThrAsnHisLys 101
              |||||||||||||||||||||||||||||||||||||
Db        145 AATGATGACTATCGTCGCGGTAAACTCACGAATCATAAG 183

RESULT 5
US-08-333-321-1
; Sequence 1, Application US/08333321
; Patent No. 5786192
;  GENERAL INFORMATION:
;    APPLICANT:  Obata, Shusei
;    TITLE OF INVENTION:  FARNESYL PYROPHOPHATE SYNTHETASE
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Edward W. Greason, Esq.
;      STREET:  1 Broadway
;      CITY:  New York
;      STATE:  N.Y.
;      COUNTRY:  U.S.A.
;      ZIP:  10004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  No. 5786192epad, Windows 3.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/333,321
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/953,424
;      FILING DATE:  29-SEP-1992
;      APPLICATION NUMBER:  JP 3-253788
;      FILING DATE:  01-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Greason, Edward W.
;      REGISTRATION NUMBER:  18,918
;      REFERENCE/DOCKET NUMBER:  077670/00310
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 425-7200
;      TELEFAX:  (212) 425-5288
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  893 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  Bacillus stearothermophilus
US-08-333-321-1

Alignment Scores:
Pred. No.:             9.13e-06    Length:        893
Score:                 14.00       Matches:       14
Percent Similarity:    100.00%     Conservative:  0
Best Local Similarity: 100.00%     Mismatches:    0
Query Match:           4.88%       Indels:        0
DB:                    1           Gaps:          0

US-09-925-637-64 (1-287) x US-08-333-321-1 (1-893)

Qy         72 GluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
              .  ||||||||||||||||||||||||||||||||||||||||
Db        226 GAAATGATCCATACGTACTCTTTGATCCATGATGATTTGCCG 267

RESULT 6
US-08-534-910B-2
; Sequence 2, Application US/08534910B
; Patent No. 5766911
;  GENERAL INFORMATION:
;    APPLICANT:  KOIKE, Ayumi
```

```
;    APPLICANT:  OBATA, Shusei
;    APPLICANT:  NISHINO, Tokuzo
;    APPLICANT:  OHNUMA, Shinichi
;    APPLICANT:  NAKAZAWA, Takeshi
;    APPLICANT:  OGURA, Kyozo
;    APPLICANT:  KOYAMA, Tanetoshi
;    TITLE OF INVENTION:  Mutated Farnesyldiphosphate Synthase Capable
;    TITLE OF INVENTION:  Of Synthesizing Geranylgeranyldiphosphate And Gene Coding Ther
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenyon & Kenyon
;      STREET:  1025 Connecticut Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  U.S.
;      ZIP:  20036-5405
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.25" Floppy Disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;      SOFTWARE:  IBM/Word Perfect 6.1 Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/534,910B
;      FILING DATE:  28-SEPT-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-25253
;      FILING DATE:  14-FEB-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Toffenetti, Judith L.
;      REGISTRATION NUMBER:  39,048
;      REFERENCE/DOCKET NUMBER:  77670/398
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)429-1776
;      TELEFAX:  (202)429-0796
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  894 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Bacillus stearothermophilus
US-08-534-910B-2

Alignment Scores:
Pred. No.:                9.14e-06        Length:         894
Score:                    14.00           Matches:        14
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              4.88%           Indels:         0
DB:                       1               Gaps:           0

US-09-925-637-64 (1-287) x US-08-534-910B-2 (1-894)

Qy         72 GluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
              ||||||||||||||||||||||||||||||||||||||||||||
Db        226 GAAATGATCCATACGTACTCTTTGATCCATGATGATTTGCCG 267

RESULT 7
US-08-534-910B-3
; Sequence 3, Application US/08534910B
; Patent No. 5766911
;  GENERAL INFORMATION:
;    APPLICANT:  KOIKE, Ayumi
;    APPLICANT:  OBATA, Shusei
;    APPLICANT:  NISHINO, Tokuzo
;    APPLICANT:  OHNUMA, Shinichi
;    APPLICANT:  NAKAZAWA, Takeshi
;    APPLICANT:  OGURA, Kyozo
;    APPLICANT:  KOYAMA, Tanetoshi
;    TITLE OF INVENTION:  Mutated Farnesyldiphosphate Synthase Capable
;    TITLE OF INVENTION:  Of Synthesizing Geranylgeranyldiphosphate And Gene Coding T
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenyon & Kenyon
;      STREET:  1025 Connecticut Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  U.S.
;      ZIP:  20036-5405
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.25" Floppy Disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;      SOFTWARE:  IBM/Word Perfect 6.1 Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/534,910B
;      FILING DATE:  28-SEPT-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-25253
;      FILING DATE:  14-FEB-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Toffenetti, Judith L.
;      REGISTRATION NUMBER:  39,048
;      REFERENCE/DOCKET NUMBER:  77670/398
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)429-1776
;      TELEFAX:  (202)429-0796
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  894 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Bacillus stearothermophilus
US-08-534-910B-3

Alignment Scores:
Pred. No.:                9.14e-06        Length:         894
Score:                    14.00           Matches:        14
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              4.88%           Indels:         0
DB:                       1               Gaps:           0

US-09-925-637-64 (1-287) x US-08-534-910B-3 (1-894)

Qy         72 GluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
              ||||||||||||||||||||||||||||||||||||||||||||
Db        226 GAAATGATCCATACGTACTCTTTGATCCATGATGATTTGCCG 267

RESULT 8
US-08-534-910B-5
; Sequence 5, Application US/08534910B
; Patent No. 5766911
;  GENERAL INFORMATION:
;    APPLICANT:  KOIKE, Ayumi
;    APPLICANT:  OBATA, Shusei
;    APPLICANT:  NISHINO, Tokuzo
;    APPLICANT:  OHNUMA, Shinichi
;    APPLICANT:  NAKAZAWA, Takeshi
;    APPLICANT:  OGURA, Kyozo
;    APPLICANT:  KOYAMA, Tanetoshi
;    TITLE OF INVENTION:  Mutated Farnesyldiphosphate Synthase Capable
;    TITLE OF INVENTION:  Of Synthesizing Geranylgeranyldiphosphate And Gene Coding T
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenyon & Kenyon
;      STREET:  1025 Connecticut Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  DC
```

```
;       COUNTRY:  U.S.
;       ZIP:  20036-5405
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.25" Floppy Disk
;       COMPUTER:  IBM PC Compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;       SOFTWARE:  IBM/Word Perfect 6.1 Windows
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/534,910B
;       FILING DATE:  28-SEPT-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  JP 7-25253
;       FILING DATE:  14-FEB-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Toffenetti, Judith L.
;       REGISTRATION NUMBER:  39,048
;       REFERENCE/DOCKET NUMBER:  77670/398
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202)429-1776
;       TELEFAX:  (202)429-0796
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  894 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
;     ORIGINAL SOURCE:
;       ORGANISM:  Bacillus stearothermophilus
US-08-534-910B-5

Alignment Scores:
Pred. No.:                9.14e-06      Length:          894
Score:                    14.00         Matches:         14
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              4.88%         Indels:          0
DB:                       1             Gaps:            0

US-09-925-637-64 (1-287) x US-08-534-910B-5 (1-894)

Qy          72 GluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
               |||||||||||||||||||||||||||||||||||||||||||
Db         226 GAAATGATCCATACGTACTCTTTGATCCATGATGATTTGCCG 267

RESULT 9
US-08-886-466-1
; Sequence 1, Application US/08886466C
; Patent No. 6040165
; GENERAL INFORMATION:
;  APPLICANT: Narita, Keishi
;  APPLICANT: Ishida, Chika
;  APPLICANT: Takeuchi, Yoshie
;  APPLICANT: Ohto, Chikara
;  APPLICANT: Ohnuma, Shinichi
;  APPLICANT: Nishino, Tokuzo
;  TITLE OF INVENTION: MUTANT PRENYL DIPHOSPHATE SYNTHASE
;  FILE REFERENCE: 77670/494
;  CURRENT APPLICATION NUMBER: US/08/886,466C
;  CURRENT FILING DATE: 1997-07-10
;  EARLIER APPLICATION NUMBER: JP 8-191635
;  EARLIER FILING DATE: 1996-07-03
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 894
;   TYPE: DNA
;   ORGANISM: Bacillus stearothermophilus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(891)
US-08-886-466-1

Alignment Scores:
Pred. No.:                9.14e-06      Length:          894
Score:                    14.00         Matches:         14
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              4.88%         Indels:          0
DB:                       3             Gaps:            0

US-09-925-637-64 (1-287) x US-08-886-466-1 (1-894)

Qy          72 GluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
               |||||||||||||||||||||||||||||||||||||||||||
Db         226 GAAATGATCCATACGTACTCTTTGATCCATGATGATTTGCCG 267

RESULT 10
US-09-475-304-1
; Sequence 1, Application US/09475304
; Patent No. 6225096
; GENERAL INFORMATION:
;  APPLICANT: Narita, Keishi
;  APPLICANT: Ishida, Chika
;  APPLICANT: Takeuchi, Yoshie
;  APPLICANT: Ohto, Chikara
;  APPLICANT: Ohnuma, Shinichi
;  APPLICANT: Nishino, Tokuzo
;  TITLE OF INVENTION: MUTANT PRENYL DIPHOSPHATE SYNTHASE
;  FILE REFERENCE: 77670/494
;  CURRENT APPLICATION NUMBER: US/09/475,304
;  CURRENT FILING DATE: 1999-12-30
;  EARLIER APPLICATION NUMBER: JP 8-191635
;  EARLIER FILING DATE: 1996-07-03
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 894
;   TYPE: DNA
;   ORGANISM: Bacillus stearothermophilus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(891)
US-09-475-304-1

Alignment Scores:
Pred. No.:                9.14e-06      Length:          894
Score:                    14.00         Matches:         14
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              4.88%         Indels:          0
DB:                       4             Gaps:            0

US-09-925-637-64 (1-287) x US-09-475-304-1 (1-894)

Qy          72 GluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
               |||||||||||||||||||||||||||||||||||||||||||
Db         226 GAAATGATCCATACGTACTCTTTGATCCATGATGATTTGCCG 267

RESULT 11
US-09-101-126-4
; Sequence 4, Application US/09101126
; Patent No. 6316216
; GENERAL INFORMATION:
;  APPLICANT: OHTO, CHIKARA
;  APPLICANT: NAKANE, HIROYUKI
;  APPLICANT: NISHINO, TOKUZO
;  APPLICANT: OHNUMA, SHINICHI
;  APPLICANT: HIROOKA, KAZUTAKE
;  TITLE OF INVENTION: MUTATED PRENYL DIPHOSPHATE SYNTHASES
;  FILE REFERENCE: 77670/566
;  CURRENT APPLICATION NUMBER: US/09/101,126
;  CURRENT FILING DATE: 1999-04-27
;  EARLIER APPLICATION NUMBER: PCT/JP97/03921
;  EARLIER FILING DATE: 1997-10-29
```

; EARLIER APPLICATION NUMBER: JP 8-307506
; EARLIER FILING DATE: 1996-11-05
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 894
; TYPE: DNA
; ORGANISM: Bacillus stearothermophilus
; FEATURE:
; OTHER INFORMATION: 256-276 is an Asp-rich coding domain
US-09-101-126-4

Alignment Scores:

Pred. No.:	9.14e-06	Length:	894
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.88%	Indels:	0
DB:	4	Gaps:	0

US-09-925-637-64 (1-287) x US-09-101-126-4 (1-894)

```
Qy         72 GluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
              |||||||||||||||||||||||||||||||||||||||||||
Db        226 GAAATGATCCATACGTACTCTTTGATCCATGATGATTTGCCG 267
```

RESULT 12
US-09-367-528A-4
; Sequence 4, Application US/09367528A
; Patent No. 6395525
; GENERAL INFORMATION:
; APPLICANT: TOYOTA JIDOSHA KABUSHIKI KAISHA
; TITLE OF INVENTION: Geranyl Diphosphate Synthetase Gene
; FILE REFERENCE: PH-586
; CURRENT APPLICATION NUMBER: US/09/367,528A
; CURRENT FILING DATE: 1999-08-16
; PRIOR APPLICATION NUMBER: JP97/346686
; PRIOR FILING DATE: 1997-12-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 894
; TYPE: DNA
; ORGANISM: Bacillus stearothermophilus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(894)
US-09-367-528A-4

Alignment Scores:

Pred. No.:	9.14e-06	Length:	894
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.88%	Indels:	0
DB:	4	Gaps:	0

US-09-925-637-64 (1-287) x US-09-367-528A-4 (1-894)

```
Qy         72 GluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
              |||||||||||||||||||||||||||||||||||||||||||
Db        226 GAAATGATCCATACGTACTCTTTGATCCATGATGATTTGCCG 267
```

RESULT 13
US-09-275-742-1
; Sequence 1, Application US/09275742
; Patent No. 6130069
; GENERAL INFORMATION:
; APPLICANT: Wilding, Edwina Imogen
; APPLICANT: Gwynn, Michael
; TITLE OF INVENTION: ispA
; FILE REFERENCE: GM10205
; CURRENT APPLICATION NUMBER: US/09/275,742
; CURRENT FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-275-742-1

Alignment Scores:

Pred. No.:	0.014	Length:	876
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.83%	Indels:	0
DB:	3	Gaps:	0

US-09-925-637-64 (1-287) x US-09-275-742-1 (1-876)

```
Qy         78 SerLeuIleHisAspAspLeuProAlaMetAsp 88
              ||||||||||||||||||||||||||||||||
Db        226 AGCTTGATTCACGATGACCTTCCTGCTATGGAT 258
```

RESULT 14
US-08-858-207A-173
; Sequence 173, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 173:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1284 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-858-207A-173

Alignment Scores:

Pred. No.:	0.0204	Length:	1284
Score:	11.00	Matches:	11

```
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            3.83%          Indels:        0
DB:                     4              Gaps:          0

US-09-925-637-64 (1-287) x US-08-858-207A-173 (1-1284)

Qy         78 SerLeuIleHisAspAspLeuProAlaMetAsp 88
              ||||||||||||||||||||||||||||||||||
Db        948 AGCTTGATTCACGATGACCTTCCTGCTATGGAT 980

RESULT 15
US-08-961-527-76/c
; Sequence 76, Application US/08961527
; Patent No. 6420135
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Streptococcus pneumoniae Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  391
     CORRESPONDENCE ADDRESS:
       ADDRESSEE:  Human Genome Sciences, Inc.
       STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,527
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB340P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  76:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10011 base pairs
       TYPE:  nucleic acid
       STRANDEDNESS:  double
       TOPOLOGY:  linear
US-08-961-527-76

Alignment Scores:
Pred. No.:              0.156          Length:        10011
Score:                  11.00          Matches:       11
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            3.83%          Indels:        0
DB:                     4              Gaps:          0

US-09-925-637-64 (1-287) x US-08-961-527-76 (1-10011)

Qy         78 SerLeuIleHisAspAspLeuProAlaMetAsp 88
              ||||||||||||||||||||||||||||||||||
Db       7653 AGCTTGATTCACGATGACCTTCCTGCTATGGAT 7621


Search completed: May 31, 2003, 00:06:30
Job time : 66 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 30, 2003, 23:42:12 ; Search time 187 Seconds
(without alignments)
2069.416 Million cell updates/sec

Title: US-09-925-637-64
Perfect score: 287
Sequence: 1 MTNLPMNKLIDEVNNELSVA.....ELTQIDEQFNTKHLLEIVDL 287

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

arched: 845702 seqs, 674182571 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1689256

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09925637/runat_23052003_174851_24218/app_query.fasta_1.455
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=oli.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09925637_@CGN_1_1_57_@runat_23052003_174851_24218
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	287	100.0	861	9	US-10-084-205-63	Sequence 63, Appl
	2	287	100.0	861	10	US-09-925-637-63	Sequence 63, Appl
c	3	287	100.0	1893	7	US-08-781-986A-155	Sequence 155, App
	4	197	68.6	882	10	US-09-815-242-8485	Sequence 8485, Ap
	5	192	66.9	864	10	US-09-815-242-4184	Sequence 4184, Ap
c	6	137	47.7	413	10	US-09-815-242-2822	Sequence 2822, Ap
c	7	112	39.0	337	10	US-09-815-242-3275	Sequence 3275, Ap
	8	18	6.3	882	10	US-09-815-242-6533	Sequence 6533, Ap
	9	18	6.3	7528	10	US-09-070-927A-55	Sequence 55, Appl
	10	15	5.2	891	9	US-09-941-947A-19	Sequence 19, Appl
	11	15	5.2	891	10	US-09-934-903-13	Sequence 13, Appl
	12	15	5.2	891	10	US-09-934-868-71	Sequence 71, Appl
	13	14	4.9	811	10	US-09-974-300-1015	Sequence 1015, Ap
	14	12	4.2	888	10	US-09-815-242-7142	Sequence 7142, Ap
	15	12	4.2	900	10	US-09-815-242-5972	Sequence 5972, Ap
	16	12	4.2	900	10	US-09-815-242-9986	Sequence 9986, Ap
	17	11	3.8	876	10	US-09-815-242-9175	Sequence 9175, Ap
	18	11	3.8	876	10	US-09-815-242-9499	Sequence 9499, Ap
	19	11	3.8	912	10	US-09-815-242-7285	Sequence 7285, Ap
	20	11	3.8	912	10	US-09-815-242-7450	Sequence 7450, Ap
	21	10	3.5	474	10	US-09-974-300-5445	Sequence 5445, Ap
	22	9	3.1	927	10	US-09-974-300-999	Sequence 999, App
c	23	9	3.1	640681	10	US-09-790-988-1	Sequence 1, Appli
	24	8	2.8	235	10	US-09-878-574-13980	Sequence 13980, A
	25	8	2.8	436	10	US-09-983-965-176	Sequence 176, App
	26	8	2.8	553	10	US-09-070-927A-587	Sequence 587, App
	27	8	2.8	586	10	US-09-864-761-12177	Sequence 12177, A
	28	8	2.8	623	9	US-10-108-915-19	Sequence 19, Appl
c	29	8	2.8	791	9	US-10-097-065-102	Sequence 102, App
	30	8	2.8	888	10	US-09-815-242-7873	Sequence 7873, Ap
	31	8	2.8	990	9	US-09-284-320-47	Sequence 47, Appl
	32	8	2.8	993	9	US-10-108-915-23	Sequence 23, Appl
	33	8	2.8	1062	9	US-10-108-915-15	Sequence 15, Appl
	34	8	2.8	1161	9	US-10-108-915-37	Sequence 37, Appl
	35	8	2.8	1215	9	US-10-066-500-38	Sequence 38, Appl
	36	8	2.8	1215	9	US-10-002-796-38	Sequence 38, Appl
	37	8	2.8	1215	9	US-10-066-273-38	Sequence 38, Appl
	38	8	2.8	1215	9	US-10-066-494-38	Sequence 38, Appl
	39	8	2.8	1215	9	US-10-066-269-38	Sequence 38, Appl
	40	8	2.8	1215	9	US-10-066-193-38	Sequence 38, Appl
	41	8	2.8	1215	9	US-10-066-211-38	Sequence 38, Appl
	42	8	2.8	1224	9	US-09-981-876-22	Sequence 22, Appl
	43	8	2.8	1224	9	US-09-148-545-22	Sequence 22, Appl
	44	8	2.8	1268	9	US-10-108-915-17	Sequence 17, Appl
	45	8	2.8	1316	9	US-09-284-320-72	Sequence 72, Appl

ALIGNMENTS

RESULT 1
US-10-084-205-63
; Sequence 63, Application US/10084205
; Publication No. US20030049648A1
; GENERAL INFORMATION:
; APPLICANT: Choi, Gil
; TITLE OF INVENTION: 37 Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB515P1
; CURRENT APPLICATION NUMBER: US/10/084,205
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/151,933
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 63
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-084-205-63

Alignment Scores:

Pred. No.:	2.02e-291	Length:	861
Score:	287.00	Matches:	287
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0

DB: 9 Gaps: 0

US-09-925-637-64 (1-287) x US-10-084-205-63 (1-861)

```
Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy         61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy         81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300

Qy        101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600

Qy        201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660

Qy        221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Qy        241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy        261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy        281 LeuLeuGluIleValAspLeu 287
              |||||||||||||||||||||
Db        841 TTATTAGAAATCGTTGATTTA 861
```

RESULT 2
US-09-925-637-63
; Sequence 63, Application US/09925637
; Patent No. US20020103338A1
; GENERAL INFORMATION:
; APPLICANT: Choi
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
; FILE REFERENCE: PB560
; CURRENT APPLICATION NUMBER: US/09/925,637
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/151,933
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 08/781,986
; PRIOR FILING DATE: 1997-01-03
; PRIOR APPLICATION NUMBER: US 08/956,171
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/009,861
; PRIOR FILING DATE: 1996-01-06
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-637-63

Alignment Scores:

Pred. No.:	2.02e-291	Length:	861
Score:	287.00	Matches:	287
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-925-637-64 (1-287) x US-09-925-637-63 (1-861)

```
Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy         61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy         81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300

Qy        101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600

Qy        201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
```

```
Qy        221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Qy        241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy        261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy        281 LeuLeuGluIleValAspLeu 287
              |||||||||||||||||||||
Db        841 TTATTAGAAATCGTTGATTTA 861

RESULT 3
US-08-781-986A-155/c
; Sequence 155, Application US/08781986A
; Publication No. US20030054436A1
  GENERAL INFORMATION:
    APPLICANT: Charles Kunsch
;   TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;   NUMBER OF SEQUENCES: 5255
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Human Genome Sciences, Inc.
;     STREET: 9410 Key West Avenue
;     CITY: Rockville
;     STATE: Maryland
;     COUNTRY: USA
;     ZIP: 20850
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;     COMPUTER: HP Vectra 486/33
;     OPERATING SYSTEM: MSDOS version 6.2
;     SOFTWARE: ASCII Text
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/781,986A
;     FILING DATE:
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:
;     FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Benson, Bob
;     REGISTRATION NUMBER: 30,446
;     REFERENCE/DOCKET NUMBER: PB248PP
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (301) 309-8504
;     TELEFAX: (301) 309-8512
  INFORMATION FOR SEQ ID NO: 155:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1893 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
US-08-781-986A-155

Alignment Scores:
Pred. No.:              4.37e-291     Length:          1893
Score:                  287.00        Matches:         287
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            100.00%       Indels:          0
DB:                     7             Gaps:            0

US-09-925-637-64 (1-287) x US-08-781-986A-155 (1-1893)

Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1314 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 1255

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1254 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 1195

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1194 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 1135

Qy         61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1134 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 1075

Qy         81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1074 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 1015

Qy        101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1014 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 955

Qy        121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 895

Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        894 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 835

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        834 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 775

Qy        181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        774 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 715

Qy        201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        714 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 655

Qy        221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 595

Qy        241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 535

Qy        261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 475

Qy        281 LeuLeuGluIleValAspLeu 287
              |||||||||||||||||||||
Db        474 TTATTAGAAATCGTTGATTTA 454

RESULT 4
US-09-815-242-8485
; Sequence 8485, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT: Ohlsen, Kari L.
;  APPLICANT: Zyskind, Judith W.
;  APPLICANT: Wall, Daniel
;  APPLICANT: Trawick, John D.
;  APPLICANT: Carr, Grant J.
;  APPLICANT: Yamamoto, Robert T.
;  APPLICANT: Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION: Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE: 2001-03-21
```

```
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8485
;   LENGTH: 882
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(882)
US-09-815-242-8485

Alignment Scores:
Pred. No.:              4.4e-197      Length:          882
Score:                  197.00        Matches:         268
Percent Similarity:     99.26%        Conservative:    0
Best Local Similarity:  99.26%        Mismatches:      1
Query Match:            68.64%        Indels:          2
DB:                     10            Gaps:            0

US-09-925-637-64 (1-287) x US-09-815-242-8485 (1-882)

Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy         61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy         81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCAT 300

Qy        101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThr-ThrLy 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||  |||||

Db        541 CTATTAACATTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGC-AACTAA 599

Qy        200 sGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLe 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 AGAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTTCAGATTAAAGATGATTT 659

Qy        220 uLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAs 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 ATTAGACTGCTATGGTGATGAAGCGAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAA 719

Qy        240 nAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTy 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 TAATAAAAGTACATACGTGAGTTTATTAGGAAAAGATGGCGCAGAAGATAAATTGACTTA 779

Qy        260 rHisArgAspAlaAlaValAspGluLeu 269
              |||||||||||||||||||||||||||
Db        780 TCATAGAGACGCAGCAGTGGATGAACTA 807

RESULT 5
US-09-815-242-4184
; Sequence 4184, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4184
;   LENGTH: 864
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-09-815-242-4184

Alignment Scores:
Pred. No.:              7.51e-192     Length:          864
Score:                  192.00        Matches:         263
Percent Similarity:     99.25%        Conservative:    0
Best Local Similarity:  99.25%        Mismatches:      1
Query Match:            66.90%        Indels:          2
DB:                     10            Gaps:            0

US-09-925-637-64 (1-287) x US-09-815-242-4184 (1-864)

Qy          6 MetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerVal 25
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTA 60

Qy         26 MetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIle 45
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATC 120

Qy         46 ArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMet 65
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATG 180

Qy         66 LysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCA 240

Qy         86 AlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGlu 105
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCATAAAGTATATGGTGAG 300

Qy        106 TrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSer 125
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGT 360

Qy        126 AspAspArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSer 145
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGT 420

Qy        146 GlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIle 165
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATT 480

Qy        166 AspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAla 185
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCACTATTAACATTTGCG 540

Qy        186 ValMetSerAlaAlaAspIleAlaAsnValAspAspThr-ThrLysGluHisLeuGluSe 205
                 ||||||||||||||||||||||||||||||||||    ||||||||||||||||||||
Db        541 GTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGC-AACTAAAGAACATTTAGAAAG 599

Qy        205 rTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGl 225
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 TTATAGTTATCATTTAGGTATGATGTTTCAGATTAAAGATGATTTATTAGACTGCTATGG 659

Qy        225 yAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTy 245
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 TGATGAAGCGAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTACATA 719

Qy        245 rValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAl 265
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 CGTGAGTTTATTAGGAAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGACGCAGC 779

Qy        265 aValAspGluLeu 269
                 |||||||||||||
          780 AGTGGATGAACTA 792

RESULT 6
US-09-815-242-2822/c
; Sequence 2822, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2822
;   LENGTH: 413
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-09-815-242-2822

Alignment Scores:
Pred. No.:                1.61e-134      Length:          413
Score:                    137.00         Matches:         137
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              47.74%         Indels:          0
DB:                       10             Gaps:            0

US-09-925-637-64 (1-287) x US-09-815-242-2822 (1-413)

Qy         54 LeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMet 73
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 TTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATG 354

Qy         74 IleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArg 93
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 ATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGA 294

Qy         94 ArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAsp 113
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 CGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGAT 234

Qy        114 AlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAspArgLeuThrAspGluVal 133
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 GCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTA 174

Qy        134 LysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGly 153
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 AAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGT 114

Qy        154 GlnMetLeuAspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIle 173
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 CAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATA 54

Qy        174 HisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAla 190
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         53 CACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCA 3

RESULT 7
US-09-815-242-3275/c
; Sequence 3275, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
```

```
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3275
;   LENGTH: 337
    TYPE: DNA
    ORGANISM: Staphylococcus aureus
   09-815-242-3275

Alignment Scores:
Pred. No.:              2.1e-108        Length:         337
Score:                  112.00          Matches:        112
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            39.02%          Indels:         0
DB:                     10              Gaps:           0

US-09-925-637-64 (1-287) x US-09-815-242-3275 (1-337)

Qy          76 ThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGly 95
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         336 ACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGA 277

Qy          96 LysLeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeu 115
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         276 AAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTA 217

Qy         116 LeuThrLysAlaPheGluLeuIleSerSerAspAspArgLeuThrAspGluValLysIle 135
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         216 TTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATA 157

Qy         136 LysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMet 155
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         156 AAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATG 97

Qy         156 LeuAspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLys 175
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          96 TTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAA 37

Qy         176 ThrLysThrGlyAlaLeuLeuThrPheAlaValMet 187
               ||||||||||||||||||||||||||||||||||||
Db          36 ACAAAAACAGGAGCATTATTAACTTTTGCGGTTATG 1

RESULT 8
US-09-815-242-6533
; Sequence 6533, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck; Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6533
;   LENGTH: 882
;   TYPE: DNA
;   ORGANISM: Enterococcus faecalis
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(882)
US-09-815-242-6533

Alignment Scores:
Pred. No.:              1.78e-09        Length:         882
Score:                  18.00           Matches:        18
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            6.27%           Indels:         0
DB:                     10              Gaps:           0

US-09-925-637-64 (1-287) x US-09-815-242-6533 (1-882)

Qy          71 LeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAsp 88
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         211 TTAGAGATGATTCATACGTATTCATTAATTCATGATGATTTACCAGCAATGGAT 264

RESULT 9
US-09-070-927A-55
; Sequence 55, Application US/09070927A
; Patent No. US20020120116A1
;   GENERAL INFORMATION:
;        APPLICANT: Charles A. Kunsch
;                   Patrick J. Dillon
;                   Steven Barash
;        TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
;        NUMBER OF SEQUENCES: 982
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/070,927A
;             FILING DATE: 04-May-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/046,655
;             FILING DATE: 1997-05-16
;             APPLICATION NUMBER: 60/044,031
;             FILING DATE: 1997-05-06
;             APPLICATION NUMBER: 60/066,009
```

```
;             FILING DATE: 1997-11-14
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Kenley K. Hoover
;             REGISTRATION NUMBER: 40,302
;             REFERENCE/DOCKET NUMBER: PB369
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (301) 309-8504
;             TELEFAX: (301) 309-8512
;    INFORMATION FOR SEQ ID NO: 55:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 7528 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 55:
US-09-070-927A-55

Alignment Scores:
Pred. No.:                1.44e-08      Length:          7528
Score:                    18.00         Matches:         18
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              6.27%         Indels:          0
DB:                       10            Gaps:            0

US-09-925-637-64 (1-287) x US-09-070-927A-55 (1-7528)

Qy          71 LeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAsp 88
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6852 TTAGAGATGATTCATACGTATTCATTAATTCATGATGATTTACCAGCAATGGAC 6905

RESULT 10
US-09-941-947A-19
; Sequence 19, Application US/09941947A
; Publication No. US20030003528A1
; GENERAL INFORMATION:
;  APPLICANT: Brzostowicz, Patricia C.
;  APPLICANT:  Cheng, Qiong
;  APPLICANT:  DiCosimo, Deana J.
;  APPLICANT:  Koffas, Mattheos
;  APPLICANT:  Miller, Edward S. Jr.
;  APPLICANT:  Odom, J. Martin
;  APPLICANT:  Picataggio, Steve
;  APPLICANT:  Rouviere, Pierre E.
;  TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
;  FILE REFERENCE: CL1903 US NA
;  CURRENT APPLICATION NUMBER: US/09/941,947A
;  CURRENT FILING DATE:  2001-09-01
;  PRIOR APPLICATION NUMBER: 60/229,907
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/229,858
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
;   LENGTH: 891
;   TYPE: DNA
;   ORGANISM: Methylomonas 16a
US-09-941-947A-19

Alignment Scores:
Pred. No.:                2.51e-06      Length:          891
Score:                    15.00         Matches:         15
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              5.23%         Indels:          0
DB:                       9             Gaps:            0

US-09-925-637-64 (1-287) x US-09-941-947A-19 (1-891)

Qy          77 TyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAsp 91
               |||||||||||||||||||||||||||||||||||||||||||||
Db        238 TATTCGCTGATTCACGACGATCTGCCGGCCATGGACAACGATGAT 282
```

```
RESULT 11
US-09-934-903-13
; Sequence 13, Application US/09934903
; Patent No. US20020102690A1
; GENERAL INFORMATION:
;  APPLICANT: Koffas, Mattheos
;  APPLICANT:  Odom, J. Martin
;  APPLICANT:  Schenzle, Andreas J.
;  APPLICANT:  No. US20020102690Alton, Kelley C.
;  APPLICANT:  Tomb, Jean-Francois
;  APPLICANT:  Rouviere, Pierre
;  APPLICANT:  Picataggio, Stephen
;  APPLICANT:  Cheng, Qiong
;  TITLE OF INVENTION: Genes Involved in Isoprenoid Compounds Production
;  FILE REFERENCE: CL1646 US NA
;  CURRENT APPLICATION NUMBER: US/09/934,903
;  CURRENT FILING DATE:  2001-08-22
;  PRIOR APPLICATION NUMBER: 60/229,907
;  PRIOR FILING DATE: September 1, 2001
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
;   LENGTH: 891
;   TYPE: DNA
;   ORGANISM: Methylomonas 16a
;   FEATURE:
;   OTHER INFORMATION: ORF7
US-09-934-903-13

Alignment Scores:
Pred. No.:                2.51e-06      Length:          891
Score:                    15.00         Matches:         15
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              5.23%         Indels:          0
DB:                       10            Gaps:            0

US-09-925-637-64 (1-287) x US-09-934-903-13 (1-891)

Qy          77 TyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAsp 91
               |||||||||||||||||||||||||||||||||||||||||||||
Db        238 TATTCGCTGATTCACGACGATCTGCCGGCCATGGACAACGATGAT 282

RESULT 12
US-09-934-868-71
; Sequence 71, Application US/09934868
; Patent No. US20020137190A1
; GENERAL INFORMATION:
;  APPLICANT: Koffas, Mattheos
;  APPLICANT:  Odom, James M
;  APPLICANT:  Schenzle, Andreas J
;  TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
;  FILE REFERENCE: CL1596 US NA
;  CURRENT APPLICATION NUMBER: US/09/934,868
;  CURRENT FILING DATE:  2001-08-22
;  PRIOR APPLICATION NUMBER: 60/229,858
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 81
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 71
;   LENGTH: 891
;   TYPE: DNA
;   ORGANISM: Methylomonas 16a
;   FEATURE:
;   OTHER INFORMATION: ORF7 ISPa
US-09-934-868-71

Alignment Scores:
Pred. No.:                2.51e-06      Length:          891
Score:                    15.00         Matches:         15
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
```

```
Query Match:            5.23%       Indels:        0
DB:                     10          Gaps:          0

US-09-925-637-64 (1-287) x US-09-934-868-71 (1-891)

Qy         77 TyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAsp 91
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        238 TATTCGCTGATTCACGACGATCTGCCGGCCATGGACAACGATGAT 282

RESULT 13
US-09-974-300-1015
; Sequence 1015, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
;  APPLICANT: Berka, Randy M.
;  APPLICANT:  Clausen, Ib Groth
;  TITLE OF INVENTION: Methods For Monitoring Multiple Gene
;  TITLE OF INVENTION:  Expression
;  FILE REFERENCE: 10085.500-US
;  CURRENT APPLICATION NUMBER: US/09/974,300
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: 09/680,598
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/279,526
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 8481
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1015
;   LENGTH: 811
;   TYPE: DNA
;   ORGANISM: Bacillus licheniformis
US-09-974-300-1015

Alignment Scores:
Pred. No.:              2.55e-05    Length:        811
Score:                  14.00       Matches:       14
Percent Similarity:     100.00%     Conservative:  0
Best Local Similarity:  100.00%     Mismatches:    0
Query Match:            4.88%       Indels:        0
DB:                     10          Gaps:          0

US-09-925-637-64 (1-287) x US-09-974-300-1015 (1-811)

Qy         72 GluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
              |||||||||||||||||||||||||||||||||||||||||||
Db        223 GAAATGATCCATACGTATTCATTAATCCATGACGACCTCCCT 264

RESULT 14
US-09-815-242-7142
; Sequence 7142, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7142
;   LENGTH: 888
;   TYPE: DNA
;   ORGANISM: Haemophilus influenzae
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(888)
US-09-815-242-7142

Alignment Scores:
Pred. No.:              0.00348     Length:        888
Score:                  12.00       Matches:       12
Percent Similarity:     100.00%     Conservative:  0
Best Local Similarity:  100.00%     Mismatches:    0
Query Match:            4.18%       Indels:        0
DB:                     10          Gaps:          0

US-09-925-637-64 (1-287) x US-09-815-242-7142 (1-888)

Qy         77 TyrSerLeuIleHisAspAspLeuProAlaMetAsp 88
              ||||||||||||||||||||||||||||||||||||
Db        238 TATTCCTTAATTCACGATGATTTACCTGCAATGGAT 273

RESULT 15
US-09-815-242-5972
; Sequence 5972, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5972
;   LENGTH: 900
;   TYPE: DNA
;   ORGANISM: Escherichia coli
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(900)
US-09-815-242-5972
```

Alignment Scores:

Pred. No.:	0.00353	Length:	900
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.18%	Indels:	0
DB:	10	Gaps:	0

US-09-925-637-64 (1-287) x US-09-815-242-5972 (1-900)

```
Qy         77 TyrSerLeuIleHisAspAspLeuProAlaMetAsp 88
              ||||||||||||||||||||||||||||||||||||
Db        235 TACTCATTAATTCATGATGATTTACCGGCAATGGAT 270
```

Search completed: May 31, 2003, 01:02:16
Job time : 192 secs

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 30, 2003, 22:32:52 ; Search time 1558 Seconds
(without alignments)
2983.379 Million cell updates/sec

Title: US-09-925-637-64
Perfect score: 287
Sequence: 1 MTNLPMNKLIDEVNNELSVA.....ELTQIDEQFNTKHLLEIVDL 287

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Word size: 1

Total number of hits satisfying chosen parameters: 32308016

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09925637/runat_23052003_174850_24121/app_query.fasta_1.455
-DB=EST -QFMT=fastap -SUFFIX=oli.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09925637_@CGN_1_1_1906_@runat_23052003_174850_24121 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	24	8.4	4484	17	BH770981	BH770981 LLMGtag70
2	10	3.5	451	17	AQ176350	AQ176350 HS_3213_A
3	10	3.5	528	17	AQ924302	AQ924302 RPCI-23-2
4	10	3.5	576	13	BI944945	BI944945 sai24d02.
c 5	10	3.5	696	17	AG051077	AG051077 Pan trogl
c 6	9	3.1	273	14	F02981	F02981 HSC1IC102 n
7	9	3.1	365	13	BM399846	BM399846 5009-0-62
8	9	3.1	422	17	BH876624	BH876624 hr32c04.b
c 9	9	3.1	474	14	N29629	N29629 yw67f05.s1
c 10	9	3.1	501	17	AQ186598	AQ186598 HS_3079_B
c 11	9	3.1	514	14	BQ616272	BQ616272 faa99h01.
12	9	3.1	516	12	BF702840	BF702840 MI-P-H2-a
13	9	3.1	546	17	BH785617	BH785617 fzmb013f0
14	9	3.1	563	14	BQ602325	BQ602325 MI-P-HO-a
15	9	3.1	570	12	BF703521	BF703521 MI-P-H1-a
c 16	9	3.1	573	13	BM095848	BM095848 fv27e04.x
17	9	3.1	617	17	BH445833	BH445833 BOHCF25TF
18	9	3.1	662	9	AL641814	AL641814 AL641814
19	9	3.1	666	13	BI717621	BI717621 1031021A0
20	9	3.1	668	13	BI726915	BI726915 1031088E0
21	9	3.1	675	13	BI723806	BI723806 1031067H0
c 22	9	3.1	685	13	BI717620	BI717620 1031021A0
23	9	3.1	685	13	BM449136	BM449136 DSA032E03
24	9	3.1	701	10	BE238154	BE238154 894040E07
25	9	3.1	722	13	BM597369	BM597369 170006874
c 26	9	3.1	783	17	BH706927	BH706927 BOMIK63TF
27	9	3.1	854	17	CNS034T9	AL227862 Tetraodon
c 28	9	3.1	855	17	AZ207859	AZ207859 SP_0135_A
29	9	3.1	1389	14	BM808194	BM808194 AGENCOURT
c 30	8	2.8	114	17	AZ849898	AZ849898 2M0151I01
31	8	2.8	128	17	BH077347	BH077347 RPCI-24-3
c 32	8	2.8	184	17	AZ874342	AZ874342 2M0188L21
33	8	2.8	197	10	BB191918	BB191918 BB191918
34	8	2.8	202	12	BE827852	BE827852 RC2-ET002
35	8	2.8	216	10	BB588852	BB588852 BB588852
36	8	2.8	243	10	AV346846	AV346846 AV346846
37	8	2.8	261	10	AV353343	AV353343 AV353343
38	8	2.8	264	9	AI846208	AI846208 UI-M-AM1-
39	8	2.8	266	10	AV343488	AV343488 AV343488
c 40	8	2.8	272	14	F02986	F02986 HSC1IE122 n
c 41	8	2.8	273	17	AZ758116	AZ758116 1M0550A02
42	8	2.8	274	14	T43543	T43543 6806 Lambda
43	8	2.8	289	9	AV146579	AV146579 AV146579
44	8	2.8	297	10	BB192476	BB192476 BB192476
45	8	2.8	300	10	AV325043	AV325043 AV325043

ALIGNMENTS

RESULT 1
BH770981
LOCUS BH770981 4484 bp DNA linear GSS 01-MAY-2002
DEFINITION LLMGtag706 MG1363 Random Sequence Tag Library Lactococcus lactis subsp. cremoris genomic, DNA sequence.
ACCESSION BH770981
VERSION BH770981.1 GI:20373938
KEYWORDS GSS.
SOURCE Lactococcus lactis subsp. cremoris.
ORGANISM Lactococcus lactis subsp. cremoris
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
REFERENCE 1 (bases 1 to 4484)
AUTHORS Bolotin,A., Ehrlich,S.D. and Sorokin,A.
TITLE Studies of genomes of dairy bacteria Lactococcus lactis

```
  JOURNAL   Sci. Aliments, (2002) In press
COMMENT     Contact: Sorokin A
            Genetique Microbienne
            INRA
            CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
            Tel: 33 1 34 65 25 16
            Fax: 33 1 34 65 25 21
            Email: sorokine@jouy.inra.fr
            best homologue in strain IL1403 is xseA (94%)
            Class: shotgun
            High quality sequence start: 30
            High quality sequence stop: 4456.
FEATURES             Location/Qualifiers
     source          1. .4484
                     /organism="Lactococcus lactis subsp. cremoris"
                     /strain="MG1363"
                     /db_xref="taxon:1359"
                     /clone_lib="MG1363 Random Sequence Tag Library"
                     /note="Vector: pSGMU2; Site_1: SmaI; Library of
                     chromosomal fragments of L.lactis strain MG1363 was
                     prepared by partial AluI digestion or by sonication."
BASE COUNT     1540 a    637 c   1023 g   1284 t
ORIGIN

Alignment Scores:
Pred. No.:                2.33e-11        Length:         4484
Score:                    24.00           Matches:        24
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              8.36%           Indels:         0
DB:                       17              Gaps:           0

US-09-925-637-64 (1-287) x BH770981 (1-4484)

Qy         78 SerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeu 97
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3867 TCGTTAATTCATGATGACTTGCCTGCGATGGACAATGATGATTATCGTCGTGGAAAGTTA 3926

Qy         98 ThrAsnHisLys 101
              ||||||||||||
Db       3927 ACGAACCATAAA 3938

RESULT 2
AQ176350
LOCUS       AQ176350                451 bp    DNA     linear   GSS 17-OCT-1998
DEFINITION  HS_3213_A1_E01_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3213 Col=1 Row=I, DNA sequence.
ACCESSION   AQ176350
VERSION     AQ176350.1  GI:3573717
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 451)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 3213  row: I  column: 1
            Class: BAC ends
            High quality sequence stop: 451.
FEATURES             Location/Qualifiers
     source          1. .451
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=3213 Col=1 Row=I"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /sex="male"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
BASE COUNT     110 a    102 c    80 g    155 t     4 others
ORIGIN

Alignment Scores:
Pred. No.:                22.7            Length:         451
Score:                    10.00           Matches:        10
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              3.48%           Indels:         0
DB:                       17              Gaps:           0

US-09-925-637-64 (1-287) x AQ176350 (1-451)

Qy         49 LeuLeuLeuLeuThrLeuAspSerLeuAsn 58
              ||||||||||||||||||||||||||||||
Db        381 CTCTTACTCCTTACTCTTGATTCACTGAAT 410

RESULT 3
AQ924302
LOCUS       AQ924302                528 bp    DNA     linear   GSS 21-DEC-1999
DEFINITION  RPCI-23-278L8.TJ RPCI-23 Mus musculus genomic clone RPCI-23-278L8,
            DNA sequence.
ACCESSION   AQ924302
VERSION     AQ924302.1  GI:6613305
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 278  row: L  column: 8
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .528
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-278L8"
                     /clone_lib="RPCI-23"
                     /sex="Female"
                     /lab_host="DH10B"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
```

selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies). "

BASE COUNT 103 a 146 c 118 g 158 t 3 others
ORIGIN

Alignment Scores:

Pred. No.:	27.8	Length:	528
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.48%	Indels:	0
DB:	17	Gaps:	0

US-09-925-637-64 (1-287) x AQ924302 (1-528)

```
Qy         48 ValLeuLeuLeuLeuThrLeuAspSerLeu 57
              ||||||||||||||||||||||||||||||
Db        405 GTATTGCTTCTCCTGACTTTGGACTCCCTA 434
```

RESULT 4
AQ944945
LOCUS BI944945 576 bp mRNA linear EST 30-NOV-2001
DEFINITION sai24d02.y1 Gm-c1053 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1053-4012 5' similar to TR:Q42698 Q42698 GERANYLGERANYL PYROPHOSPHATE SYNTHASE ;, mRNA sequence.
ACCESSION BI944945
VERSION BI944945.1 GI:16282306
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 576)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann ,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 391.

FEATURES Location/Qualifiers
source 1. .576
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1053-4012"
/clone_lib="Gm-c1053"
/tissue_type="Whole seedling, 3 week old, greenhouse grown"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2: XhoI; The Harosoy NIL was constructed and seed was provided by Dr. J. Specht, University of Nebraska (Shoemaker and Specht, 1995). The cDNA library was constructed from mRNA isolated from whole seedlings of 3 week old greenhouse grown plants. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site and a 3' anchor. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in cooperation with Dr. Paul Keim's laboratory at Northern Arizona University."

BASE COUNT 112 a 210 c 144 g 110 t
ORIGIN

Alignment Scores:

Pred. No.:	31.2	Length:	576
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.48%	Indels:	0
DB:	13	Gaps:	0

US-09-925-637-64 (1-287) x BI944945 (1-576)

```
Qy         40 AlaGlyGlyLysArgIleArgProValLeu 49
              |||||||||||||||||||||||||||||||
Db        265 GCCGGCGGAAAGAGGATCCGCCCGGTGCTC 294
```

RESULT 5
AG051077/c
LOCUS AG051077 696 bp DNA linear GSS 02-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-031L05.F, genomic survey sequence.
ACCESSION AG051077
VERSION AG051077.1 GI:16587969
KEYWORDS GSS.
SOURCE Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male BAC Library clone:PTB-031L05.F.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE BAC end sequences of Library PTB
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 696)
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.

FEATURES Location/Qualifiers
source 1. .696
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="PTB-031L05.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

BASE COUNT 208 a 133 c 118 g 237 t
ORIGIN

Alignment Scores:

Pred. No.:	39.8	Length:	696

```
Score:                  10.00          Matches:       10
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            3.48%          Indels:        0
DB:                     17             Gaps:          0

US-09-925-637-64 (1-287) x AG051077 (1-696)

Qy        276 PheAsnThrLysHisLeuLeuGluIleVal 285
              ||||||||||||||||||||||||||||||
Db        201 TTTAACACCAAGCATTTATTGGAAATAGTA 172

RESULT 6
F02981/c
LOCUS       F02981                 273 bp    mRNA    linear   EST 02-FEB-1995
DEFINITION  HSC1IC102 normalized infant brain cDNA Homo sapiens cDNA clone
            c-1ic10 3', mRNA sequence.
ACCESSION   F02981
VERSION     F02981.1  GI:646538
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 273)
  AUTHORS   Auffray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes
            ,M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F.,
            Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y.,
            Sebastiani-Kabaktchis,C. and Tessier,A.
  TITLE     IMAGE: molecular integration of the analysis of the human genome
            and its expression
  JOURNAL   C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
  MEDLINE   95277534
COMMENT     Contact: Genethon
            Genexpress-Genethon
            Genethon Centre de recherche sur le Genome Humain
            1,rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
            Tel: 33169472800
            Fax: 33160778698
            Email: genexpress@genethon.fr
            Single read.  removed at sequence 5'end
            Genexpress_library_idt: C; Genexpress_sequence_idt: a1c-1ic10
            Seq primer: (-21)M13_universal.
FEATURES              Location/Qualifiers
     source           1. .273
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="c-1ic10"
                      /clone_lib="normalized infant brain cDNA"
                      /sex="Female"
                      /tissue_type="total brain"
                      /dev_stage="3 months old"
                      /note="Organ: brain; Vector: lafmid BA; Site_1: HindIII;
                      Site_2: NotI; sex=Female; dev_stage=3 months old;
                      isolate=muscular atrophy patient;  tissue_type=total brain
                      ;  total mRNA was oligo-(dT) primed and directionally
                      cloned 5' -> 3' into the  HindIII -> NotI sites of the
                      lafmid BA vector. Clone library from B.Soares,  Psychiatry
                      Dept. Columbia University, USA.  Normalization_method:
                      Bento Soares, P.N.A.S in press"
BASE COUNT       88 a     69 c     45 g     69 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              105            Length:        273
Score:                  9.00           Matches:       9
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            3.14%          Indels:        0
DB:                     14             Gaps:          0

US-09-925-637-64 (1-287) x F02981 (1-273)

Qy         51 LeuLeuThrLeuAspSerLeuAsnThr 59
              |||||||||||||||||||||||||||
Db        233 TTGCTAACGCTTGACTCACTTAACACT 207

RESULT 7
BM399846
LOCUS       BM399846                 365 bp    mRNA    linear   EST 17-JAN-2002
DEFINITION  5009-0-62-D06.t.1 Chilcoat/Turkewitz cDNA (large fraction)
            Tetrahymena thermophila cDNA, mRNA sequence.
ACCESSION   BM399846
VERSION     BM399846.1  GI:18199899
KEYWORDS    EST.
SOURCE      Tetrahymena thermophila.
  ORGANISM  Tetrahymena thermophila
            Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;
            Hymenostomatida; Tetrahymenina; Tetrahymena.
REFERENCE   1  (bases 1 to 365)
  AUTHORS   Turkewitz,A.P., Karrer,K.M., Jahn,C., Orias,E., Kirk,K.E., Frankel
            ,J. and Klobutcher,L.
  TITLE     EST from Tetrahymena thermophila, strain CU428.1, growing cells
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Turkewitz AP
            Molecular Genetics and Cell Biology
            University of Chicago
            920 E. 58th Street, Chicago, IL 60637, USA
            Tel: 773 702 4374
            Fax: 773 702 3172
            Email: apturkew@midway.uchicago.edu
            Seq primer: T3.
FEATURES              Location/Qualifiers
     source           1. .365
                      /organism="Tetrahymena thermophila"
                      /strain="CU428.1"
                      /db_xref="taxon:5911"
                      /clone_lib="Chilcoat/Turkewitz cDNA (large fraction)"
                      /note="Vector: BlueScript2 SK+; Details on library
                      preparation can be found in Chilcoat and Turkewitz (2001)
                      Proc. Natl. Acad. Sci USA, 98: 8709-8713."
BASE COUNT      110 a     75 c     81 g     97 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              153            Length:        365
Score:                  9.00           Matches:       9
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            3.14%          Indels:        0
DB:                     13             Gaps:          0

US-09-925-637-64 (1-287) x BM399846 (1-365)

Qy         48 ValLeuLeuLeuLeuThrLeuAspSer 56
              ||||||||||||||||||||||||||||
Db        291 GTGTTGCTGTTGCTAACACTAGACTCC 317

RESULT 8
BH876624
LOCUS       BH876624                 422 bp    DNA     linear   GSS 05-AUG-2002
DEFINITION  hr32c04.b1 WGS-ZmaysF (JM107 adapted methyl filtered) Zea mays
            genomic clone hr32c04 5', DNA sequence.
ACCESSION   BH876624
VERSION     BH876624.1  GI:22112521
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 422)
  AUTHORS   Rabinowicz,P.D., O'Shaughnessy,A.L., Balija,V., Dedhia,N.,
            Katzenburger,F., King,L., Miller,B., Muller,S., Nascimento,L.,
            Zutavern,T., McCombie,W.R. and Martienssen,R.A.
  TITLE     Genomic shotgun sequences from Zea mays (methyl-filtered)
```

```
  JOURNAL   Unpublished (2002)
COMMENT     Contact: W. Richard McCombie
            Lita Annenberg Hazen Genome Sequencing Center
            Cold Spring Harbor Laboratory
            PO Box 100, Cold Spring Harbor, NY 11724, USA
            Tel: 516 367 8884
            Fax: 516 367 8874
            Email: mccombie@cshl.org
            Plate: hr32  row: c  column: 04
            Seq primer: -21M13UnivFwd
            Class: shotgun
            High quality sequence stop: 422.
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="Zea mays"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="hr32c04"
                     /clone_lib="WGS-ZmaysF (JM107 adapted methyl filtered)"
                     /lab_host="JM107 or DH5a"
                     /note="Organ: immature ears; Site_1: Xba I; Site_2: Xba I;
                     The vector was digested with XbaI and one nucleotide was
                     added by  fill in in the recessive 3' end. The genomic DNA
                     was nebulized, end  repaired, adaptor ligated and size
                     fractionated using sephadex. The  resulting fragments were
                     between 0.8 and 3 kb and were cloned into  the vector
                     (.x/y reads in M13mp19, .b/g reads in pUC19). The same
                     ligation was transformed in either JM107 or DH5a. "
BASE COUNT      127 a     66 c     82 g    147 t
ORIGIN

Alignment Scores:
Pred. No.:                185          Length:         422
Score:                    9.00         Matches:        9
Percent Similarity:       100.00%      Conservative:   0
Best Local Similarity:    100.00%      Mismatches:     0
Query Match:              3.14%        Indels:         0
DB:                       17           Gaps:           0

US-09-925-637-64 (1-287) x BH876624 (1-422)

Qy         241 AsnLysSerThrTyrValSerLeuLeu 249
               |||||||||||||||||||||||||||
Db         162 AATAAGTCGACCTATGTGTCGTTGCTT 188

RESULT 9
N29629/c
LOCUS       N29629                 474 bp    mRNA    linear   EST 05-JAN-1996
DEFINITION  yw67f05.s1 Soares_placenta_8to9weeks_2NbHP8to9W Homo sapiens cDNA
            clone IMAGE:257313 3', mRNA sequence.
ACCESSION   N29629
VERSION     N29629.1  GI:1148149
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 474)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 405
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: m13 -40 forward
            High quality sequence stop: 405.
FEATURES             Location/Qualifiers
     source          1. .474
                     /organism="Homo sapiens"
                     /db_xref="GDB:3886923"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:257313"
                     /clone_lib="Soares_placenta_8to9weeks_2NbHP8to9W"
                     /dev_stage="two placentae: one from 8 weeks and another
                     from 9 weeks post conception"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGATTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo."
BASE COUNT      131 a    120 c    119 g     99 t      5 others
ORIGIN

Alignment Scores:
Pred. No.:                215          Length:         474
Score:                    9.00         Matches:        9
Percent Similarity:       100.00%      Conservative:   0
Best Local Similarity:    100.00%      Mismatches:     0
Query Match:              3.14%        Indels:         0
DB:                       14           Gaps:           0

US-09-925-637-64 (1-287) x N29629 (1-474)

Qy          51 LeuLeuThrLeuAspSerLeuAsnThr 59
               |||||||||||||||||||||||||||
Db         225 TTGCTAACGCTTGACTCACTTAACACT 199

RESULT 10
AQ186598/c
LOCUS       AQ186598               501 bp    DNA     linear   GSS 01-NOV-1998
DEFINITION  HS_3079_B1_B09_T7 CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3079 Col=17 Row=D, DNA sequence.
ACCESSION   AQ186598
VERSION     AQ186598.1  GI:3586040
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 501)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 3079  row: D  column: 17
            Class: BAC ends
            High quality sequence stop: 501.
FEATURES             Location/Qualifiers
     source          1. .501
                     /organism="Homo sapiens"
```

```
                     /db_xref="taxon:9606"
                     /clone="Plate=3079 Col=17 Row=D"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /sex="male"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
BASE COUNT      120 a    111 c     93 g    175 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:                231           Length:          501
Score:                    9.00          Matches:         9
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              3.14%         Indels:          0
DB:                       17            Gaps:            0

US-09-925-637-64 (1-287) x AQ186598 (1-501)

          235 GlySerAspLeuGluAsnAsnLysSer 243
              |||||||||||||||||||||||||||
          420 GGGAGTGATCTCGAAAATAATAAGAGC 394

RESULT 11
BQ616272/c
LOCUS       BQ616272                514 bp    mRNA    linear   EST 26-JUN-2002
DEFINITION  faa99h01.x1 Sugano SJD adult male Danio rerio cDNA clone 6034033
            3', mRNA sequence.
ACCESSION   BQ616272
VERSION     BQ616272.1  GI:21605941
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 514)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
            ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
            ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
            Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
            Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
            and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA
            Sequencing by: Washington University Genome Sequencing Center Clone
            distribution information can be found through the I.M.A.G.E.
            Consortium/LLNL, send email to: info@image.llnl.gov
            zebrafish identity (p-value greater than 1e-99) found to:
            gi|2594822|gb|AA658668|AA658668 fa56f02.s1 zebrafish gridded kidney
            Danio
            Seq primer: T7 from Gibco
            High quality sequence stop: 513.
FEATURES             Location/Qualifiers
     source          1. .514
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="6034033"
                     /clone_lib="Sugano SJD adult male"
                     /sex="male"
                     /tissue_type="whole body"
                     /dev_stage="adult"
                     /lab_host="DH10B (phage resistant)"
                     /note="Vector: pME18S-FL3; Site_1: DraIII (CACCATGTG);
                     Site_2: DraIII (CACTGTGTG); 1st strand cDNA was primed
                     with an oligo(dT) primer  [ATGTGGCCTTTTTTTTTTTTTTTT];
                     double-stranded cDNA was ligated to a  DraIII adaptor
                     [TGTTGGCCTACTGG], digested and cloned into distinct DraIII
                     sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site
                     CACCATGTG). XhoI should be used to isolate the cDNA
                     insert. Size selection was performed to exclude fragments
                     <1.5kb. Library constructed and donated by Dr. Sumio
                     Sugano (University of Tokyo Institute of Medical Science).
                     Custom primers for sequencing: 5' end primer
                     CTTCTGCTCTAAAAGCTGCG and 3' end primer
                     CGACCTGCAGCTCGAGCACA. "
BASE COUNT      152 a     98 c     137 g     127 t
ORIGIN

Alignment Scores:
Pred. No.:                239           Length:          514
Score:                    9.00          Matches:         9
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              3.14%         Indels:          0
DB:                       14            Gaps:            0

US-09-925-637-64 (1-287) x BQ616272 (1-514)

Qy           46 ArgProValLeuLeuLeuLeuThrLeu 54
                ||||||||||||||||||||||||||
Db          254 AGACCGGTTCTCCTGCTCCTTACACTT 228

RESULT 12
BF702840
LOCUS       BF702840                516 bp    mRNA    linear   EST 22-DEC-2000
DEFINITION  MI-P-H2-abh-d-02-1-UM.s1 MI-P-H2 Sus scrofa cDNA clone
            MI-P-H2-abh-d-02-1-UM 3', mRNA sequence.
ACCESSION   BF702840
VERSION     BF702840.1  GI:11988248
KEYWORDS    EST.
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 516)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
COMMENT     Contact: Tuggle CK
            Molecular Genetics Laboratory, Department of Animal Science
            Iowa State University
            201 Kildee Hall, Ames, IA 50011-3150, USA
            Tel: 5152944252
            Fax: 5152942401
            Email: cktuggle@iastate.edu
            The sequence contained an oligo-dT track that was present in the
            oligonucleotide that was used to prime the synthesis of first
            strand cDNA and therefore this may represent a bonafide poly A
            tail. The sequence tag present in the cDNA between the NotI site
            and the oligo-dT track served to verify it as a clone from the
            non-normalized hypothalamus at estrus day 5 library cDNA Library
            Preparation: RJ Woods, JA Green, RS Prather S142 Animal Science
            Research Center, Department of Animal Science, University of
            Missouri-Columbia, 65211 Clone distribution: clones will be
            available through Research Genetics (www.resgen.com) The following
            repetitive elements were found in this cDNA sequence: 1-57,
            >AT_rich#Low_complexity 213-248, >AT_rich#Low_complexity
            Seq primer: M13 Forward
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .516
                     /organism="Sus scrofa"
                     /strain="crossbreed"
                     /db_xref="taxon:9823"
                     /clone="MI-P-H2-abh-d-02-1-UM"
                     /clone_lib="MI-P-H2"
```

/lab_host="DH10B (Life Technologies)"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: EcoRI; The MI-P-H2 library is derived from hypothalamus at estrus day 5. For a detailed description of the library from which this clone was derived, please visit our web site at http://pigest.genome.iastate.edu/.
TAG_LIB=MI-P-H2
TAG_TISSUE=hypothalamus at estrus day 5
TAG_SEQ=TCTTCG"

BASE COUNT 156 a 84 c 82 g 193 t 1 others
ORIGIN

Alignment Scores:

Pred. No.:	240	Length:	516
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.14%	Indels:	0
DB:	12	Gaps:	0

US-09-925-637-64 (1-287) x BF702840 (1-516)

```
Qy        198 ThrThrLysGluHisLeuGluSerTyr 206
              ||||||||||||||||||||||||||
Db        449 ACAACAAAAGAACACCTAGAGTCATAC 475
```

RESULT 13
BH785617
LOCUS BH785617 546 bp DNA linear GSS 28-MAR-2002
DEFINITION fzmb013f045e05k0 fzmb filtered library Zea mays genomic clone fzmb013f045e05 5', DNA sequence.
ACCESSION BH785617
VERSION BH785617.1 GI:19790466
KEYWORDS GSS.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 546)
AUTHORS Budiman,M.A., Freese,R.G., Bedell,J.A., Nunberg,A.N. and Lakey,N.D.
TITLE GeneThresher methylation filtered genomic sequences from maize
JOURNAL Unpublished (2002)
COMMENT Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: fzmb013f045 row: e column: 05
Seq primer: SK reverse
Class: shotgun
High quality sequence stop: 546.
FEATURES Location/Qualifiers
source 1. .546
/organism="Zea mays"
/cultivar="MO17"
/db_xref="taxon:4577"
/clone="fzmb013f045e05"
/clone_lib="fzmb filtered library"
/note="Organ: leaf; Vector: pBCSK(-); Site_1: HincII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincII-digested pBCSK(-) vector and electroporated into E. coli cells."
BASE COUNT 141 a 102 c 131 g 172 t
ORIGIN

Alignment Scores:

Pred. No.:	258	Length:	546
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.14%	Indels:	0
DB:	17	Gaps:	0

US-09-925-637-64 (1-287) x BH785617 (1-546)

```
Qy        241 AsnLysSerThrTyrValSerLeuLeu 249
              ||||||||||||||||||||||||||
Db         72 AATAAGTCGACCTATGTGTCGTTGCTT 98
```

RESULT 14
BQ602325
LOCUS BQ602325 563 bp mRNA linear EST 24-JUN-2002
DEFINITION MI-P-HO-ahy-c-10-1-UM.s1 MI-P-HO Sus scrofa cDNA clone MI-P-HO-ahy-c-10-1-UM 3', mRNA sequence.
ACCESSION BQ602325
VERSION BQ602325.1 GI:21549051
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 563)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Tuggle CK
Molecular Genetics Laboratory, Department of Animal Science
Iowa State University
201 Kildee Hall, Ames, IA 50011-3150, USA
Tel: 5152944252
Fax: 5152942401
Email: cktuggle@iastate.edu
Tissue Procurement: Dr. Chris Tuggle, Iowa State University
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA sequence: 29-59, >AT_rich#Low_complexity (matched compliment) 215-251, >AT_rich#Low_complexity (matched compliment)
Seq primer: M13 FORWARD
POLYA=Yes.
FEATURES Location/Qualifiers
source 1. .563
/organism="Sus scrofa"
/strain="crossbreed"
/db_xref="taxon:9823"
/clone="MI-P-HO-ahy-c-10-1-UM"
/clone_lib="MI-P-HO"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: EcoRI; The MI-P-HO library is a normalized library derived from hypothalamus at estrus days 0 and 12 and ovary at estrus days 0, 5 and 12. For a detailed description of the library from which this clone was derived, please visit our web site at http://pigest.genome.iastate.edu/.
TAG_LIB=MI-P-HO
TAG_TISSUE=d_0_hypothalamus
TAG_SEQ=TAGATG"
BASE COUNT 177 a 94 c 91 g 201 t
ORIGIN

Alignment Scores:

Pred. No.:	269	Length:	563
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.14%	Indels:	0

DB: 14 Gaps: 0

US-09-925-637-64 (1-287) x BQ602325 (1-563)

```
Qy        198 ThrThrLysGluHisLeuGluSerTyr 206
              ||||||||||||||||||||||||||
Db        452 ACAACAAAAGAACACCTAGAGTCATAC 478
```

RESULT 15
BF703521
LOCUS BF703521 570 bp mRNA linear EST 22-DEC-2000
DEFINITION MI-P-H1-abm-c-10-1-UM.s1 MI-P-H1 Sus scrofa cDNA clone MI-P-H1-abm-c-10-1-UM 3', mRNA sequence.
ACCESSION BF703521
VERSION BF703521.1 GI:11988929
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 570)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Tuggle CK
Molecular Genetics Laboratory, Department of Animal Science
Iowa State University
201 Kildee Hall, Ames, IA 50011-3150, USA
Tel: 5152944252
Fax: 5152942401
Email: cktuggle@iastate.edu
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the non-normalized hypothalamus at estrus day 0 library cDNA Library Preparation: RJ Woods, JA Green, RS Prather S142 Animal Science Research Center, Department of Animal Science, University of Missouri-Columbia, 65211 Clone distribution: clones will be available through Research Genetics (www.resgen.com) The following repetitive elements were found in this cDNA sequence: 1-59, >AT_rich#Low_complexity 215-252, >AT_rich#Low_complexity
Seq primer: M13 Forward
POLYA=Yes.
FEATURES Location/Qualifiers
source 1. .570
/organism="Sus scrofa"
/strain="crossbreed"
/db_xref="taxon:9823"
/clone="MI-P-H1-abm-c-10-1-UM"
/clone_lib="MI-P-H1"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: EcoRI; The MI-P-H1 library is derived from hypothalamus at estrus day 0. For a detailed description of the library from which this clone was derived, please visit our web site at http://pigest.genome.iastate.edu/.
TAG_LIB=MI-P-H1
TAG_TISSUE=hypothalamus at estrus day 0
TAG_SEQ=TAGATG"
BASE COUNT 179 a 95 c 93 g 203 t
ORIGIN

Alignment Scores:

Pred. No.:	273	Length:	570
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.14%	Indels:	0

DB: 12 Gaps: 0

US-09-925-637-64 (1-287) x BF703521 (1-570)

```
Qy        198 ThrThrLysGluHisLeuGluSerTyr 206
              ||||||||||||||||||||||||||
Db        453 ACAACAAAAGAACACCTAGAGTCATAC 479
```

Search completed: May 31, 2003, 00:05:14
Job time : 1566 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 30, 2003, 13:08:18 ; Search time 2483 Seconds
(without alignments)
10091.629 Million cell updates/sec

Title: US-09-925-637-63
Perfect score: 861
Sequence: 1 atgacgaatctaccgatgaa.....tattagaaatcgttgattta 861

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	859.4	99.8	333750	1	AP004827	AP004827 Staphyloc
c	2	854.6	99.3	301550	1	AP003134	AP003134 Staphyloc
c	3	854.6	99.3	346900	1	AP003362	AP003362 Staphyloc
	4	853	99.1	882	6	AR106466	AR106466 Sequence
	5	449.8	52.2	3444	1	AF270104	AF270104 Staphyloc
	6	449.8	52.2	3444	6	AX145422	AX145422 Sequence
c	7	449.8	52.2	4045	1	AF269889	AF269889 Staphyloc
c	8	449.8	52.2	4045	6	AX145207	AX145207 Sequence
	9	406.8	47.2	783	6	AX141501	AX141501 Sequence
	10	252.6	29.3	250050	1	AL591978	AL591978 Listeria
	11	239.8	27.9	5635	6	AX416832	AX416832 Sequence
	12	233.4	27.1	1080	1	AB003187	AB003187 Micrococc
c	13	231.2	26.9	195269	6	AX417035	AX417035 Sequence
	14	231.2	26.9	333050	1	AL596168	AL596168 Listeria
	15	231.2	26.9	349980	6	AX417044	AX417044 Sequence
	16	222.6	25.9	2268	1	LMO012349	AJ012349 Listeria
c	17	218.6	25.4	300950	1	AP001516	AP001516 Bacillus
c	18	218.4	25.4	11055	1	AE010637	AE010637 Fusobacte
c	19	214.6	24.9	12353	1	AE007710	AE007710 Clostridi
c	20	204	23.7	3454	1	AF269958	AF269958 Staphyloc
c	21	204	23.7	3454	6	AX145276	AX145276 Sequence
c	22	202.4	23.5	296750	1	AP003191	AP003191 Clostridi
c	23	182	21.1	135599	8	CPU30821	U30821 Cyanophora
c	24	178.4	20.7	13536	1	AE010066	AE010066 Streptoco
c	25	178.4	20.7	50463	1	AE014159	AE014159 Streptoco
c	26	175.2	20.3	10165	1	AE006583	AE006583 Streptoco
	27	174.8	20.3	10029	1	AE006320	AE006320 Lactococc
	28	164	19.0	811	6	AX432600	AX432600 Sequence
	29	163.2	19.0	348050	1	AP003581	AP003581 Nostoc sp
c	30	161.2	18.7	218470	1	BSUB0013	Z99116 Bacillus su
	31	161.2	18.7	282700	1	BACJH642	D84432 Bacillus su
	32	157.2	18.3	894	6	E11869	E11869 Mutated DNA
	33	155.8	18.1	1337	6	AX021147	AX021147 Sequence
	34	155.6	18.1	894	6	AR148386	AR148386 Sequence
	35	155.6	18.1	894	6	E05337	E05337 DNA encodin
	36	155.6	18.1	894	6	E11866	E11866 Mutated DNA
	37	155.6	18.1	894	6	E11867	E11867 Mutated DNA
	38	155.6	18.1	894	6	E11868	E11868 Mutated DNA
	39	155.6	18.1	894	6	E14762	E14762 gDNA encodi
	40	155.6	18.1	894	6	E16095	E16095 DNA encodin
	41	155.6	18.1	894	6	E27521	E27521 Geranyl dip
	42	155.6	18.1	894	6	E54864	E54864 Process for
	43	155.6	18.1	1260	1	BACFDPS	D13293 B. stearoth
	44	154	17.9	894	6	E11865	E11865 Mutated DNA
	45	154	17.9	894	6	E27520	E27520 Geranyl dip

ALIGNMENTS

RESULT 1
AP004827/c
LOCUS AP004827 333750 bp DNA linear BCT 02-JUL-2002
DEFINITION Staphylococcus aureus subsp. aureus MW2 DNA, complete genome, strain:MW2, section 6/10.
ACCESSION AP004827 BA000033
VERSION AP004827.1 GI:21204509
KEYWORDS .
SOURCE Staphylococcus aureus subsp. aureus MW2 (strain:MW2) DNA.
ORGANISM Staphylococcus aureus subsp. aureus MW2
Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE 1
AUTHORS Baba,T., Takeuchi,F., Kuroda,M., Yuzawa,H., Aoki,K., Oguchi,A., Nagai,Y., Iwama,N., Asano,K., Naimi,T., Kuroda,H., Cui,L., Yamamoto,K. and Hiramatsu,K.

```
  TITLE     Genome and virulence determinants of high virulence
            community-acquired MRSA
  JOURNAL   Lancet 359 (9320), 1819-1827 (2002)
  MEDLINE   22040717
   PUBMED   12044378
REFERENCE   2  (bases 1 to 333750)
  AUTHORS   Aoki,K., Oguchi,A., Nagai,Y., Asano,K., Iwama,N., Baba,T.,
            Kuroda,M., Hiramatsu,K. and Kikuchi,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAR-2002) Akio Oguchi, National Institute of
            Technology and Evaluation, Biotechnology Center; 2Chome 49-10
            Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
            (E-mail:oguchi@nite.go.jp, URL:http://www.bio.nite.go.jp/,
            Tel:81-3-3481-8423, Fax:81-3-3481-8424)
FEATURES             Location/Qualifiers
     source          1. .333750
                     /organism="Staphylococcus aureus subsp. aureus MW2"
                     /strain="MW2"
                     /db_xref="taxon:196620"
     gene            complement(278. .619)
                     /gene="MW1341"
     CDS             complement(278. .619)
                     /gene="MW1341"
                     /note="ORFID:MW1341"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAB95206.1"
                     /db_xref="GI:21204510"
                     /translation="MTNITKEVFDNLEQEIDVFAKNKTLGSSEAKPYLDEYHSKIIDY
                     FKQVNDITGNIDFDKLNQYPVVPMNFKERYDYMIERKYHFMGYRQMKTFKTELIKMNA
                     SYQTRLKNKQV"
     gene            complement(624. .1283)
                     /gene="nth"
     CDS             complement(624. .1283)
                     /gene="nth"
                     /note="ORFID:MW1342"
                     /codon_start=1
                     /transl_table=11
                     /product="endonuclease-like protein"
                     /protein_id="BAB95207.1"
                     /db_xref="GI:21204511"
                     /translation="MVSKKKALEMIDVIANMFPDAECELKHDNPFELTIAVLLSAQCT
                     DVLVNRVTTELFKKYKTPEDYLAVSDEELMNDIRSIGLYRNKAKNIKKLCQSLIDQFN
                     GEIPQTHKELESLAGVGRKTANVVMSVAFDEPSLAVDTHVERVSKRLGINRWKDNVRQ
                     VEDRLCSVIPRDRWNRSHHQLIFFGRYHCLARKPKCDICPLLEDCREGQKRYKASLKE
                     A"
     gene            complement(1273. .1959)
                     /gene="MW1343"
     CDS             complement(1273. .1959)
                     /gene="MW1343"
                     /note="ORFID:MW1343
                     hypothetical protein, similar to chromosome replication
                     initiation protein dnaD"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB95208.1"
                     /db_xref="GI:21204512"
                     /translation="MDKYQLKARPVVIRRELLDHYSDLGLDEQDLVILLKLIYASETS
                     NKQPSIELLQKGSTMQPRDITMVIQNLIQRELLELQVQKDEEGRFTEYMNLDPFFEKL
                     SHILKQQSMETKEQNSKEKFKQLFRVLEDTFARPLSPYEIETLNQWIDVDKHDSAIIQ
                     AALDEANSLNKLSFKYMDRILLNWKKNNVKTIDDSRKIREKFNKPKMTHTVKTVPKFD
                     WLNGENLDGK"
     gene            complement(2287. .3579)
                     /gene="asnS"
     CDS             complement(2287. .3579)
                     /gene="asnS"
                     /note="ORFID:MW1344"
                     /codon_start=1
                     /transl_table=11
                     /product="asparaginyl-tRNA synthetase"
                     /protein_id="BAB95209.1"
                     /db_xref="GI:21204513"
                     /translation="MKTTIKQAKDHLNQDVTIGAWLTNKRSSGKIAFLQLRDGTGFMQ
                     GVVVKSEVDEEVFKLAKEITQESSLYVTGTITEDNRSDLGYEMQVKSIEVISEAHDYP
                     ITPKNHGTEFLMDHRHLWLRSKKQHAVMKIRNEVIRATYEFFNKDGFTKVDPPILTAS
                     APEGTSELFHTKYFDQDAFLSQSGQLYLEAAAMAHGKVFSFGPTFRAEKSKTRRHLIE
                     FWMIEGEMAFTNHAESLEIQEQYVTHVVKSVLENCKLELKILERDTSKLEKVATPFPR
                     ISYDDAIEFLKSEGFDDIEWGEDFGAPHETAIANHYDLPVFITNYPTKIKPFYMQPNP
                     ENEETVLCADLIAPEGYGEIIGGSERVDDLELLEQRVKEHGLDEEAYSYYLDLRRYGS
                     VPHCGFGLGLERTVAWISGVEHVRETAPFPRLLNRLYP"
     gene            complement(3901. .6594)
                     /gene="dinG"
     CDS             complement(3901. .6594)
                     /gene="dinG"
                     /note="ORFID:MW1345"
                     /codon_start=1
                     /transl_table=11
                     /product="probable ATP-dependent DNA helicase dinG"
                     /protein_id="BAB95210.1"
                     /db_xref="GI:21204514"
                     /translation="MGMATYAVVDLETTGNQLDFDDIIQIGITFVRNNQIIDTYHSMI
                     RTNLEIPPFIQALTSIEENMLQQAPYFNQVAQEIYDKIKDCIFVAHNVDFDLNFIKKA
                     FKDCNIQYRPKKVIDTLEIFKIAFPTDKSYQLSELAEAHGITLANAHRADEDAATTAK
                     LMILAFEKFEKLPLDTLKQLYYLSKQLKYDLYDIFFEMVRQYDAKPLDKSYEKFEQII
                     YRKQVDFKKPTTNYNGSLKSLYSKAVDQLGLTYRPQQLYLAETILDQLMHSEKAMIEA
                     SLGSGKSLAYLLAALMYNIETGKHVMISTNTKLLQSQLLEKDIPAMNEALNFKINALL
                     IKSKSDYISLGLISQILKDDTSNYEVNILKMQLLIWITETPSGDIQELNLKGGQKMYF
                     DQKIETYVPARHDVHYYNFIKRNAQNIQIGITNHAHLIHSDVENSIYQLFDDCIVDEA
                     HRLPDYALNQVTNELSYADIKYQLGLIGKNENEKLLKAIDQLEKQRILEKLDIAPIDI
                     FGLKASMNEIHELNEQLFSTIFTIINDSDVYDDDIHRFHNVFTFETKDILKDLHAIID
                     KLNKTLEIFNGISHKTVKSLRKQLLYLKDKFKNIEQSLKAGHTSFISIKNLSQKSTIR
                     LYVKDYAVKDVLTKQVLEKFKSLIFISGTLKFNHSFEAFKQLFNKDVHFNTFEVNTSL
                     QSAKNTSVFIPSDVASYQYKNIDEYVASIVSYIIEYTTITSSKCLVLFTSYKMMHMVQ
                     DMLNELPEFEDYVVLTQQQNQNYKIVQQFNNFDKAILLGTSTFFEGFDFQANGIKCVM
                     IAKLPFMNKHNAKYWLMDSEFTSTFKEYVLPDAVTRFRQGLGRLIRSENDRGIIVSFD
                     DRLINSNYKNFFEQTLENYRQKKGDIQQFGKLLRQIQKKKK"
     gene            complement(6618. .7589)
                     /gene="MW1346"
     CDS             complement(6618. .7589)
                     /gene="MW1346"
                     /note="ORFID:MW1346"
                     /codon_start=1
                     /transl_table=11
                     /product="truncated hypothetical protein"
                     /protein_id="BAB95211.1"
                     /db_xref="GI:21204515"
                     /translation="MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLK
                     LEGCKIDSVNHKGHLLQQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSL
                     VGNQSSFFILSDEQTKGRGRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALG
                     IRDAIQHFSQDEVKVKWPNDIYIDNGKVCGFLTEMVANNDGIEAIICGIGINLTQQLE
                     NFDESIRHRATSVQLHDKNKLDRYQFLERLLQEIEKRYNQFLTLPFSEIREEYIAASN
                     IWNRTLLFTENDKQFKGQAIDLDYDGYLIVRDEAGESHRLISADIDF"
     gene            complement(7576. .8778)
                     /gene="MW1347"
     CDS             complement(7576. .8778)
                     /gene="MW1347"
                     /note="ORFID:MW1347
                     hypothetical protein, similar to poly(A) polymerase"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB95212.1"
                     /db_xref="GI:21204516"
                     /translation="MDKSLFEQARPILEQIQDNGFEAYYVGGSVRDYVMGRNIHDIDI
                     TTSATPDEIESIFSHTIPVGKEHGTINVVFNDENYEVTTFRAEEDYVDHRRPSGVTFV
                     RDLYEDLQRRDFTMNAIAMDTAYKLYDYFDGQQDINNRIIRTVGIAEERFQEDALRMI
                     RCLRFQSQLSFDIATETFEAMRIQMADIKFLSIERIVIELTKLMRGINVEKSFNHLKS
                     LKAFNYMPYFEHLDMNQINVTEAIDLELLIAIVSVKFDINYSLKPLKLSNRQVKDINQ
                     YIQIMNALPSIITKEQLKMFVYDYDTHLIKNVMVAADVIKANDIQGHEPLIVNLQTID
                     ETLHRLPMHNRKDMMVNGGVLMAHLNAKSGPWLKDVLRQIEIAIVTGKVSNEETEILK
                     WVDNHVKI"
     gene            complement(8783. .9925)
                     /gene="MW1348"
     CDS             complement(8783. .9925)
                     /gene="MW1348"
                     /note="ORFID:MW1348
```

```
                     hypothetical protein, similar to lipopolysaccharide
                     biosynthesis-related pr homolog"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB95213.1"
                     /db_xref="GI:21204517"
                     /translation="MKIGITCYPSMGGSGIIATELGIKLAERGHEVHFITSNIPFRIR
                     KPLPNMIFHQVEVNQYAVFQYPPYDITLSTKIAEVIKEYGLDLLHMHYAVPHAICGIL
                     AREMSGKDIKIMTTLHGTDITVLGYDHSLQGAIKFGIEKSDIVTSVSKSLAQETHEII
                     ETNKEIIPIYNFVRENEFPTKHNTALKSQFGIAPDEKVLIHVSNFRQVKRIDTIIETF
                     AKVREKIPSKLILLGDGPELVPMRQLTKELNVEEDVLFLGKQDCVSEFYQLSDLVLLL
                     SEKESFGLTLLEAMKTGVVPIGSNAGGIKEVIKHGETGFVVDVGDCDSASDYAIRLLE
                     DKVLYNKLQKNMLADIAERFGSELITDQYEYYYQKMLNEHNKSKGE"
     gene            complement(10167. .10484)
                     /gene="MW1349"
     CDS             complement(10167. .10484)
                     /gene="MW1349"
                     /note="ORFID:MW1349"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB95214.1"
                     /db_xref="GI:21204518"
                     /translation="MKSMVEMQREVDEYIGQFKTGYFSPLANLARLTEEVGELAREIN
                     HTYGEKKKKDSEEANTIKAELGDNLFVLLCLANSMGIDMTESFNETMEKFNTRDKNRF
                     ERK"
     gene            complement(10820. .11518)
                     /gene="MW1350"
     CDS             complement(10820. .11518)
                     /gene="MW1350"
                     /note="ORFID:MW1350"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB95215.1"
                     /db_xref="GI:21204519"
                     /translation="MSSIFTMIIYFVILMVLPLWAQHKVKSNYEKYSQVRSTSGKTGR
                     EVALEILHANGIYDVEVVKGEGFLTDHYDPKKKVVSLSPANYDRPSVAGTAIAAHEVG
                     HAIQDHQGYWFLRFRAALVPVANLGSSLSYMIIMLGIILTAIGSAFGSTALWIGAGLM
                     SLAVLFSIVTLPVEFDASSRAMKQITALNIVNEKEYKHARKVLSAAAMTYVAATAVAV
                     AELVRIILLARSSD"

Query Match             99.8%;  Score 859.4;  DB 1;  Length 333750;
Best Local Similarity   99.9%;  Pred. No. 4.7e-154;
Matches  860;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118374 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 118315

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118314 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 118255

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118254 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 118195

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118194 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 118135

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118134 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 118075

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118074 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 118015

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118014 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 117955

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117954 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 117895

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117894 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 117835

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db     117834 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAA 117775

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117774 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 117715

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117714 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 117655

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117654 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 117595

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117594 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 117535

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||||||||||||||||||||
Db     117534 TTATTAGAAATCGTTGATTTA 117514


RESULT 2
AP003134/c
LOCUS       AP003134               301550 bp    DNA     linear   BCT 02-JUL-2002
DEFINITION  Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
            genome, section 6/10.
ACCESSION   AP003134 BA000018
VERSION     AP003134.2  GI:14349226
KEYWORDS    .
SOURCE      Staphylococcus aureus subsp. aureus N315 (sub_species:aureus N315,
            strain:N315) DNA.
  ORGANISM  Staphylococcus aureus subsp. aureus N315
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
            Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M.,
            Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
            Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
            Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
            Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
            Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
  TITLE     Whole genome sequencing of meticillin-resistant Staphylococcus
            aureus
  JOURNAL   Lancet 357 (9264), 1225-1240 (2001)
  MEDLINE   21311952
   PUBMED   11418146
REFERENCE   2  (bases 1 to 301550)
  AUTHORS   Aoki,K., Oguchi,A., Hosoyama,A., Nagai,Y., Kuroda,M., Hiramatsu,K.
            and Kikuchi,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JAN-2001) Akio Oguchi, National Institute of
            Technology and Evaluation, Biotechnology Center; 2Chome 49-10
            Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
            (E-mail:oguchi@nite.go.jp, URL:http://www.bio.nite.go.jp/,
            Tel:81-3-3481-8423, Fax:81-3-3481-8424)
COMMENT     On Jun 12, 2001 this sequence version replaced gi:13701258.
FEATURES             Location/Qualifiers
     source          1. .301550
                     /organism="Staphylococcus aureus subsp. aureus N315"
                     /strain="N315"
```

```
                /sub_species="aureus N315"
                /db_xref="taxon:158879"
     gene       complement(249. .947)
                /gene="SA1293"
     CDS        complement(249. .947)
                /gene="SA1293"
                /note="ORFID:SA1293"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="BAB42553.1"
                /db_xref="GI:13701259"
                /translation="MSSIFTMIIYFVILMVLPLWAQHKVKSNYEKYSQVRSTSGKTGR
                EVALEILHANGIYDVEVVKGEGFLTDHYDPKKKVVSLSPANYDRPSVAGTAIAAHEVG
                HAIQDHQGYWFLRFRAALVPVANLGSSLSYMIIMLGIILTAIGSAFGSTALWIGAGLM
                SLAVLFSIVTLPVEFDASSRAMKQITALNIVNEKEYKHARKVLSAAAMTYVAATAVAV
                AELVRIILIARSSD"
     gene       complement(1001. .1588)
                /gene="SA1294"
     CDS        complement(1001. .1588)
                /gene="SA1294"
                /note="ORFID:SA1294"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="BAB42554.1"
                /db_xref="GI:13701260"
                /translation="MTINTFWQYTLYQRSWLMMLLICNILGMIYGYIWYGEQLSHTPW
                HFKIFVPDSPTAILYLVISISLILIQKQNSIIEALAFVTLFKYGIWAVIMNILFIIEQ
                GDITVNGLVLMFSHSIMAVQAIYFYPRFKRSMIGISVAMIWVFLNDYIDYFHLQFPYY
                DFITTHVWQIGVLSCCLSVFGLLLYIELNKLLKCK"
     gene       complement(1578. .2153)
                /gene="SA1295"
     CDS        complement(1578. .2153)
                /gene="SA1295"
                /note="ORFID:SA1295"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="BAB42555.1"
                /db_xref="GI:13701261"
                /translation="MSETLNQIKESFIEYLLFQYRFKSRIAVWVLNYIKVNEAKLANI
                HFVDTKINHHETLEIAEVGSHASAIQFTKRNIKLMNTNEIFDYIANHNCAFDIQIHFA
                NVSKREQRLDDLIVAQLTESPSYQTYLHDLNSMAIDRHKHALLIDYLLHNIDLSLQMN
                EKQRFYQLTQILNTLKLVNKHNQFEDLADDD"
     gene       complement(2167. .3411)
                /gene="SA1296"
     CDS        complement(2167. .3411)
                /gene="SA1296"
                /note="ORFID:SA1296"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="BAB42556.1"
                /db_xref="GI:13701262"
                /translation="MEDIYKLIDDINLQKLENLDSRVNEAITTDNDDALFILGETLYN
                FGLMPQGLEVFRVLYHKYPDESELLIYFIEGLMSENQTDEALEYLSYVEPSPEKLMLE
                ADLYQQINMMEVAIDKLQEALELEPNDPIIHFALAEMLYYDGQYLRATSEYETVLETG
                EYQVNGVNLFSRMADCSLQSGNYSDAIRLYDEINEDEMTSEDYLKKAISYDKNDITQE
                AIKIMTTLLSKDPDYIQGYLYLQSLYENEKNYPDAIETGKEGLRLSQFYKELMYTTGC
                LEIEHGDANEGVLLLKQALEVDNAYQEPLLILSDLYRNEEDYEAIIELLTYVDEEDLD
                PTFMWHLAFAYGQEERDKEAQHFFELAYPTMKTNIDFMSDYYFYLIEIGQKEQAITIL
                NQLLELEPSNENWHDELSRLQS"
     gene       complement(3418. .4716)
                /gene="aroA"
     CDS        complement(3418. .4716)
                /gene="aroA"
                /note="ORFID:SA1297"
                /codon_start=1
                /transl_table=11
                /product="3-phosphoshikimate 1-carboxyvinyltransferase"
                /protein_id="BAB42557.1"
                /db_xref="GI:13701263"
                /translation="MVSEQIIDISGPLKGEIEVPGDKSMTHRAIMLASLAEGTSNIYK
                PLLGEDCRRTMDIFRLLGVDIKEDEDKLVVNSPGYKAFKTPHQVLYTGNSGTTTRLLA
                GLLSGLGIESVLSGDVSIGKRPMDRVLRPLKLMDANIEGIEDNYTPLIIKPSVIKGIN
                YQMEVASAQVKSAILFASLFSNDTTVIKELDVSRNHTETMFRHFNIPIEAERLSITTT
                PDAIQHIKPADFHVPGDISSAAFFIVAALITPESDVTIHNVGINPTRSGIIDIVEKMG
                GNIQLFNQTTGAEPTASIRIQYTPMLQPITIEGELVPKAIDELPVIALLCTQAVGTST
                IKDAEELKVKETNRIDTTADMLNLLGFELQPTNDGLIIHPSEFKTNATVDSLTDHRIG
                MMLAVASLLSSEPVKIKQFDAVNVSFPGFLPKLKLLENEG"
     gene       complement(4726. .5790)
                /gene="aroB"
     CDS        complement(4726. .5790)
                /gene="aroB"
                /note="ORFID:SA1298"
                /codon_start=1
                /transl_table=11
                /product="3-dehydroquinate synthase"
                /protein_id="BAB42558.1"
                /db_xref="GI:13701264"
                /translation="MKLQTTYPSNNYPIFVEHGAIDHISTYIDQFDQSFILIDEHVNQ
                YFADKFDDILSYENVHKVIIPAGEKTKTFEQYQETLEYILSHHVTRNTAIIAVGGGAT
                GDFAGFVAATLLRGVHFIQVPTTILAHDSSVGGKVGINSKQGKNLIGAFYRPTAVIYD
                LDFLKTLPFEQILSGYAEVYKHALLNGESTTQEIEQHFKDREILQSLNGMDKYIAKGI
                ETKLDIVVADEKEQGVRKFLNLGHTFGHAVEYNHKIAHGHAVMIGIIYQFIVANILFN
                SNHDIQHYINYLTKLGYPLETITDIDFETIYQYMLSDKKNDKQGVQMVLIKHFGDIVV
                QHIDQTTLQHACEQLKTYFK"
     gene       complement(5816. .6982)
                /gene="aroC"
     CDS        complement(5816. .6982)
                /gene="aroC"
                /note="ORFID:SA1299"
                /codon_start=1
                /transl_table=11
                /product="chorismate synthase"
                /protein_id="BAB42559.1"
                /db_xref="GI:13701265"
                /translation="MRYLTSGESHGPQLTVIVEGIPANLEIKVEDINKEMFKRQGGYG
                RGRRMQIEKDTVEIVSGVRNGYTLGSPITMVVTNDDFTHWRKIMGAAPISEEERENMK
                RTITKPRPGHADLVGGMKYNHRDLRNVLERSSARETAARVAVGALCKVLLQQLDIDIY
                SRVVEIGGIKDKDFYDSETFKANLDRNDVRVIDDSIAQAMRDKIDEAKNEGDSIGGVV
                QVVVENMPVGVGSYVHYDRKLDGKIAQGVVSINAFKGVSFGEGFKAAEKPGSEIQDEI
                LYNSEIGYYRGSNHLGGLEGGMSNGMPIIVNGVMKPIPTLYKPLNSVDINTKEDFKAT
                IERSDSCAVPAASIVCEHVVAFEIAKALLEEFQSNHIEQLQQQIADRRQLNVEF"
     gene       complement(7368. .7568)
                /gene="SA1300"
     CDS        complement(7368. .7568)
                /gene="SA1300"
                /note="ORFID:SA1300"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB42560.1"
                /db_xref="GI:13701266"
                /translation="MKNIISIILGILMFLKLMELLYGAIFLDKPLNPITKIIFILTLI
                YIFYVLVKELIIFLKSKYNKSA"
     gene       complement(7597. .7722)
                /gene="SAS045"
     CDS        complement(7597. .7722)
                /gene="SAS045"
                /note="ORFID:SAS045"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB42561.1"
                /db_xref="GI:13701267"
                /translation="MWLVIMIIHNKARFLCCSKLLEKFYSFLVTRIKKRGIFFIK"
     gene       complement(7777. .8226)
                /gene="ndk"
     CDS        complement(7777. .8226)
                /gene="ndk"
                /note="ORFID:SA1301"
                /codon_start=1
                /transl_table=11
                /product="nucleoside diphosphate kinase"
```

```
                     /protein_id="BAB42562.1"
                     /db_xref="GI:13701268"
                     /translation="MERTFLMIKPDAVQRNLIGEVISRIERKGLKLVGGKLMQVPMEL
                     AETHYGEHQGKPFYNDLISFITSAPVFAMVVEGEDAVNVSRHIIGSTNPSEASPGSIR
                     GDLGLTVGRNIIHGSDSLESAEREINLWFNENEITSYASPRDAWLYE"
     gene            complement(8318. .9277)
                     /gene="gerCC"
     CDS             complement(8318. .9277)
                     /gene="gerCC"
                     /note="ORFID:SA1302"
                     /codon_start=1
                     /transl_table=11
                     /product="heptaprenyl diphosphate syntase component II"
                     /protein_id="BAB42563.1"
                     /db_xref="GI:13701269"
                     /translation="MAKLNMNNEIKKVEQRLEKAIKSKDSVLEQASLHLLSSGGKRVR
                     PAFVILSSQFGKDEQTSEQTYQVAVALELIHMATLVHDDVIDKSDKRRGKLTISKKWD

  Query Match           99.3%;  Score 854.6;  DB 1;  Length 301550;
  Best Local Similarity  99.5%;  Pred. No. 3.9e-153;
  Matches  857;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    64130 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 64071

Qy       61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
            |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    64070 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGCATGTTGTATTCATTAAATGCT 64011

Qy      121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    64010 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 63951

Qy      181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
            |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    63950 TATGAGTTAGGTGTGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 63891

Qy      241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
            |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    63890 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCAT 63831

Qy      301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    63830 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 63771

Qy      361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    63770 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 63711

Qy      421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    63710 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 63651

Qy      481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    63650 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 63591

Qy      541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
            |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    63590 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAA 63531

Qy      601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    63530 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 63471

Qy      661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    63470 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 63411

Qy      721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    63410 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 63351

Qy      781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    63350 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 63291

Qy      841 TTATTAGAAATCGTTGATTTA 861
            |||||||||||||||||||||
Db    63290 TTATTAGAAATCGTTGATTTA 63270


RESULT 3
AP003362/c
LOCUS       AP003362               346900 bp    DNA     linear   BCT 07-FEB-2002
DEFINITION  Staphylococcus aureus subsp. aureus Mu50 genomic DNA, complete
            sequence, section 5/9.
ACCESSION   AP003362 BA000017
VERSION     AP003362.2  GI:14247083
KEYWORDS    .
SOURCE      Staphylococcus aureus subsp. aureus Mu50 (sub_species:aureus Mu50,
            strain:Mu50) DNA.
  ORGANISM  Staphylococcus aureus subsp. aureus Mu50
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
            Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M.,
            Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
            Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
            Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
            Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
            Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
  TITLE     Whole genome sequencing of meticillin-resistant Staphylococcus
            aureus
  JOURNAL   Lancet 357 (9264), 1225-1240 (2001)
  MEDLINE   21311952
REFERENCE   2  (bases 1 to 346900)
  AUTHORS   Ohta,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Toshiko Ohta, University of Tsukuba College
            of Medical Technology and Nursing, Department of Medical
            Technology; 1-1-1 Ten-noodai, Tsukuba, Ibaraki 305-8577, Japan
            (E-mail:tohta@sakura.cc.tsukuba.ac.jp, Tel:81-298-53-3454,
            Fax:81-298-53-3454)
COMMENT     On May 29, 2001 this sequence version replaced gi:13875626.
FEATURES             Location/Qualifiers
     source          1. .346900
                     /organism="Staphylococcus aureus subsp. aureus Mu50"
                     /strain="Mu50"
                     /sub_species="aureus Mu50"
                     /db_xref="taxon:158878"
     gene            12. .263
                     /gene="SAV1313"
     CDS             12. .263
                     /gene="SAV1313"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAB57475.1"
                     /db_xref="GI:14247084"
                     /translation="MLTLIKLKEDEQVIIYEYIPEDDISNGKGSVTFNKKDAEVIDFS
                     LSEIENEEYFMLYRNKSFSVVRDFIEKQEFPENYKIAWY"
     gene            987. .1181
                     /gene="SAV1314"
     CDS             987. .1181
                     /gene="SAV1314"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAB57476.1"
                     /db_xref="GI:14247085"
                     /translation="MNLQSYIGKLVKLTLTNNKILIGKVIDFDDKVDNFDGYNSIEID
                     TGRITYDISENKIKTIVLQK"
     gene            complement(1725. .2750)
                     /gene="SAV1315"
```

```
CDS             complement(1725. .2750)
                /gene="SAV1315"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="BAB57477.1"
                /db_xref="GI:14247086"
                /translation="MISFENDYLEGAHEKVLKRLVDTNLVQASGYGFDQFTAQAIEKI
                KDTIDCPNATIRFLVGGTQTNQVVINSMLESYEGVISADTGHVAVHEGGAIEYSGHKV
                ITIPSKEGKVSASDVETYMETFKSDFKKDHMVFPGMVYISHPTEYGTLYSKSELEELC
                KVCKQYQLPLFMDGARLGYGLMSDQSDMTIKDIAKYCDVFYIGGTKIGALCGEAIVFT
                KNNEPKQFTTRIKHHGALLAKGRLTGIQFLELFTDNLYFNISRHAIEMANKMKDGFKN
                KGYRLYFDSPTNQQFFILSNEKIAELEQKVKFAVWEKYDDQHRVVRFATSWATTEENL
                NKLLELI"
gene            3020. .3217
                /gene="SAV1316"
CDS             3020. .3217
                /gene="SAV1316"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB57478.1"
                /db_xref="GI:14247087"
                /translation="MEQIKLKTFTAETLELLEKNINAFLSSEEATNLKLVNITIKEIE
                ERTFPNNEEEFNAILTLSVNK"
gene            3238. .4755
                /gene="SAV1317"
CDS             3238. .4755
                /gene="SAV1317"
                /codon_start=1
                /transl_table=11
                /product="cardiolipin synthetase homolog"
                /protein_id="BAB57479.1"
                /db_xref="GI:14247088"
                /translation="MFFVPFFIGEKTMRYTFSNDLGTLFTIILAIGFIINLVLAFIII
                FLERNRRTASSTWAWLFVLFVLPLIGFILYLFFGRTVSARKLNKNNGNVLTDFDGLLK
                QQIESFDKGNYGTDNKQVQKHHDLVSMLLMDQDGFLTENNKVDHFIDGNDLYDQVLKD
                IKNAKEYIHLEYYTFALDGLGKRILHALEEKLKQGLEVKILYDDVGSKNVKMANFDHF
                KSLGGEVEAFFASKLPLLNFRMNNRNHRKIIVIDGQLGYVGGFNIGDEYLGLGKLGYW
                RDTHLRIQGDAVDALQLRFILDWNSQAHRPQFEYDVKYFPKKNGPLGNSPIQIAASGP
                ASDWHQIEYGYTKMIMSAKKSVYLQSPYFIPDNSYINAIKIAAKSGVDVHLMIPCKPD
                HPLVYWATFSNASDLLSSGVKIYTYENGFIHSKMCLIDDEIVSVGTANMDFRSFELNF
                EVNAFVYDENLAKDLRVAYEHDITKSKQLTKESYANRPLSVKFKESLAKLVSPIL"
gene            4939. .5436
                /gene="SAV1318"
CDS             4939. .5436
                /gene="SAV1318"
                /note="ATP-binding protein"
                /codon_start=1
                /transl_table=11
                /product="ABC transporter homolog"
                /protein_id="BAB57480.1"
                /db_xref="GI:14247089"
                /translation="MIQISNINKSFNKRCVLKNISFDIEQGKCIALIGKNGAGKSTLI
                DILIGNVNANSGEIFDKDKLLQSENRSIMFQKTMFPDQLKVIEIINLYQSFYENPLPL
                EEIIELTKFDSSQLNQFVNKLSGGQQRLLDFVLSLIGQPQLILLDEPTSTMDIEIREY
                FWSII"
gene            5461. .5838
                /gene="SAV1319"
CDS             5461. .5838
                /gene="SAV1319"
                /note="ATP-binding protein"
                /codon_start=1
                /transl_table=11
                /product="ABC transporter homolog"
                /protein_id="BAB57481.1"
                /db_xref="GI:14247090"
                /translation="MLYTSHYIEEVERMSDKIILIENGEIILNDSTSHIRTNQQSQIT
                LSDEYIRKLKLDKDDLVIQKNHNGTIKIITSNVNDTILYLQQLHINLDDIEIQKVSIV
                DSYFNNKKQRGSNYDTKLLENRI"
gene            5807. .6538
                /gene="SAV1320"
CDS             5807. .6538
                /gene="SAV1320"
                /note="similar to ABC transporter integral membrane
                protein"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB57482.1"
                /db_xref="GI:14247091"
                /translation="MILSYLKIEFKVIMRKKTTLILSILFPVIFYILFTSILELPEDV
                KPKFYKEYMYSMTVYSLLSFSLLTFPLDIINEKQNEWRQRLMVTPFTFTSYYISKVVK
                TMLQFAIAILVIFMVGHFYKGVAMSAVQWLESGIFLWLGASLLITFGILFSLLNDIQK
                TSALANIVTIGLAVLGGLWFPINTFPNWLQHVAHVLPSYHLRKLGVDIASNHHINLIS
                FAIILLYVLGSIIAVYCISHFKRAE"
gene            6542. .7633
                /gene="SAV1321"
CDS             6542. .7633
                /gene="SAV1321"
                /note="similar to two-component sensor histidine kinase"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB57483.1"
                /db_xref="GI:14247092"
                /translation="MKFLKDTSIAEISSILYLIFPIAGIFFNEVYGPKWLYIISVIVF
                SLSYLILVIVNNRLNTLMFYILLIIHYFIICYFVFSVHPMLSLFFFYSAFAVPFTFKN
                NVKKTATNLFILTMIICTIITYLLYNNYFVAMMVYYYVISLIMLDNFKKMKNREYQKE
                IAEKNRHINTLIAEQERHRIGQDLHDTLGHVFASLSLKSELAYKLIDTDVEKVKAELL
                AINKLSRESLNKVREIIDDVKLPSFIEEIDSIRKVLKDADIDFTFENKELAQVLSPTK
                QSMLVMITREAINNVIKHANASKVHGKLKTVNNHKLLLMIEDDGKGIDSDCEVKSISQ
                RVQHLNGTLAVDSTNGTKIIIEISTGGIA"
gene            7630. .8232
                /gene="SAV1322"
CDS             7630. .8232
                /gene="SAV1322"
                /note="similar to two-component response regulator"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB57484.1"
                /db_xref="GI:14247093"
                /translation="MTSLIIAEDQNMLRQAMVQLIKLHGDFEILADTDNGLDAMKLIE
                EYNPNVVILDIEMPGMTGLEVLAEIRKKHLNIKVIIVTTFKRPGYFEKAVVNDVDAYV
                LKERSIEELVETINKVNNGEKEYSATLMTSFFVDKNPLTPKEQIVLREIGNGLSSKEI
                SEKLFLTDGTVRNYTSVIIDKLFADNRFDAWKKANEKGWI"
gene            complement(8345. .8533)
                /gene="SAV1323"
CDS             complement(8345. .8533)
                /gene="SAV1323"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB57485.1"
                /db_xref="GI:14247094"
                /translation="MNISNENTVTLIAVIIAIIIGIFLQIFFKLPLIVTAVLSILLGI
                FVGFIVYLIVSFYNKRKN"
gene            8672. .9205
                /gene="nuc"
CDS             8672. .9205
                /gene="nuc"
                /note="SAV1324"
                /codon_start=1
                /transl_table=11
                /product="thermonuclease"
                /protein_id="BAB57486.1"
                /db_xref="GI:14247095"
                /translation="MKSNKSLAMIVVAIIIVGVLAFQFMNHTGPFKKGTNHETVQDLN
```

Query Match 99.3%; Score 854.6; DB 1; Length 346900;
Best Local Similarity 99.5%; Pred. No. 3.8e-153;
Matches 857; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     251159 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 251100

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db     251099 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGCATGTTGTATTCATTAAATGCT 251040

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251039 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 250980

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db     250979 TATGAGTTAGGTGTGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 250920

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db     250919 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCAT 250860

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250859 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 250800

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250799 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 250740

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250739 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 250680

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250679 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 250620

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db     250619 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAA 250560

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250559 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 250500

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250499 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 250440

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250439 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 250380

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250379 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 250320

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||||||||||||||||||||
Db     250319 TTATTAGAAATCGTTGATTTA 250299


RESULT 4
AR106466
LOCUS       AR106466                882 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 1 from patent US 6107058.
ACCESSION   AR106466
VERSION     AR106466.1  GI:12820996
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 882)
  AUTHORS   Gwynn,M. and Wilding,E.Imogen.
  TITLE     ispA from Staphylococcus aureus
  JOURNAL   Patent: US 6107058-A 1 22-AUG-2000;

FEATURES             Location/Qualifiers
     source          1. .882
                     /organism="unknown"
BASE COUNT      328 a    118 c    186 g    250 t
ORIGIN

  Query Match          99.1%;  Score 853;  DB 6;  Length 882;
  Best Local Similarity   99.4%;  Pred. No. 2.2e-152;
  Matches  856;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGCATGTTGTATTCATTAAATGCT 120

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTGTGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              |||||||||||||||||||||||||||||||||||||| | ||||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAAGGAAATTAACAAATCAT 300

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAA 600

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||||||||||||||||||||
Db        841 TTATTAGAAATCGTTGATTTA 861


RESULT 5
AF270104
```

```
LOCUS       AF270104               3444 bp    DNA     linear   BCT 01-AUG-2000
DEFINITION  Staphylococcus epidermidis strain SR1 clone step.1049b08 genomic
            sequence.
ACCESSION   AF270104
VERSION     AF270104.1  GI:9624010
KEYWORDS    .
SOURCE      Staphylococcus epidermidis.
  ORGANISM  Staphylococcus epidermidis
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1  (bases 1 to 3444)
  AUTHORS   Kimmerly,W.J., Taylor,J.David., Nelsen,A.J., Godlevski,M.M.,
            Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
            Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,
            Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
            Furdon,P.J.
  TITLE     Transposon-mediated sequencing of the Staphylococcus epidermidis
            genome
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3444)
  AUTHORS   Taylor,J.David., Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M.,
            Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
            Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,
            Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
            Furdon,P.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAY-2000) Departments of Genomic Sciences and
            Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
            Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES             Location/Qualifiers
     source          1. .3444
                     /organism="Staphylococcus epidermidis"
                     /strain="SR1"
                     /db_xref="taxon:1282"
                     /clone="step.1049b08"
BASE COUNT     1365 a    451 c    571 g   1057 t
ORIGIN
```

Query Match 52.2%; Score 449.8; DB 1; Length 3444;
Best Local Similarity 70.2%; Pred. No. 6.4e-76;
Matches 604; Conservative 0; Mismatches 257; Indels 0; Gaps 0;

```
Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              |||| ||| |||| ||||||||||||||||| ||  | | ||||    | |        |
Db        766 ATGAAGAAACTACAGATGAATAAATTAATAAATATAATTAATACATCACTGAATAAGTCA 825

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||| |   ||||  | |  | ||| |  ||||||||||||||   ||||||||||||||
          826 ATACAATCATCACCATTAAAAACTAATTTAGAAGAAAGTATGAAATATTCATTAAATGCT 885

          121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              || |||||| | ||| |||||||  | |||||| | ||  || |    || || |  ||
Db        886 GGTGGTAAAAGAATCAGACCAGTCATATTATTATTAACACTAAAAATGCTTAACAAAGAT 945

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              ||| |   |||  | || || ||  | ||| | ||||||||||||||| |||||  | |||
Db        946 TATCAACAAGGACTAAATAGTGCTTTAGCATTGGAAATGATTCATACTTATTCTTTAATT 1005

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              ||||||||  |||||||| |||||||||||| ||||| ||  ||||||||||||||||||||
Db       1006 CATGATGATTTACCAGCAATGGATAATGACGATTACCGTAGAGGAAAATTAACAAATCAT 1065

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||| |||||||| ||||  || ||  | || |||||||| |||||||| ||||| |||
Db       1066 AAAGTTTATGGTGAATGGAAAGCCATTCTTGCTGGTGATGCATTATTAACAAAAGCTTTT 1125

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||| |  |||| | ||||  ||    |    |||   || ||| |||   || || || |
Db       1126 GAATTAGTTTCTAATGATACTACCATTGAAGATAGTGTGAAAGTAAGTATTATAAAAAGA 1185

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              || |||| |||||||| ||| | |||||||| || || |||  | | |||||| ||||
Db       1186 CTTTCAAAAGCAAGTGGACATTTGGGAATGGTGGGTGGCCAAGCGCTTGATATGGAAAGT 1245

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              |||||  |  |||||  | | |||||||| |||   || ||  |||| || ||||| ||
Db       1246 GAAGGGAAGTCAATTCGTTTAGAAACTTTAGAATCAATTCATGAAACTAAGACAGGCGCT 1305

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              ||| ||| |||| | ||||||  ||| | ||| || ||| | || ||  | |  | |
Db       1306 TTACTAAATTTTTCAGTTATGGCTGCGGTAGACATTGCTCAAGTAGAACAAAATATTGCT 1365

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              |  |||||||    | |||| |||||||||| |||||||| || ||||||||||||||
Db       1366 AAGAATTTAGATGAATTTAGTCATCATTTAGGAATGATGTTTCAAATTAAAGATGATTTA 1425

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              | ||    ||||||||||||| |||| | || |||||||| ||||| ||| | | ||||
Db       1426 CTGGATGTGTATGGTGATGAATCAAAACTTGGCAAAAAAGTAGGCAGTGATATAGTAAAT 1485

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              |||||||||| || ||   |||| | || ||||| || |||||||| || || |   ||
Db       1486 CATAAAAGTACTTATGTTTCTTTACTTGGAAAAGAAGGAGCAGAAGAAAAGTTAAACAAT 1545

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              |||  | |    ||  || |    |||  ||||||  ||| ||||   ||||     |
Db       1546 CATCAATATCTTGCTATGAACTGCTTAAATCAAATTTCTGATCAATATGATACTTCTGAA 1605

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||   || || || ||||||
Db       1606 TTAAGTGATATTGTAGATTTA 1626
```

```
RESULT 6
AX145422
LOCUS       AX145422               3444 bp    DNA     linear   PAT 31-MAY-2001
DEFINITION  Sequence 4144 from Patent WO0134809.
ACCESSION   AX145422
VERSION     AX145422.1  GI:14283987
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 3444)
  AUTHORS   Kimmerly,W.J.
  TITLE     Staphylococcus epidermidis nucleic acids and proteins
  JOURNAL   Patent: WO 0134809-A 4144 17-MAY-2001;
            GLAXO GROUP LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .3444
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic nucleic acid sequence"
BASE COUNT     1365 a    451 c    571 g   1057 t
ORIGIN
```

Query Match 52.2%; Score 449.8; DB 6; Length 3444;
Best Local Similarity 70.2%; Pred. No. 6.4e-76;
Matches 604; Conservative 0; Mismatches 257; Indels 0; Gaps 0;

```
Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              |||| ||| |||| ||||||||||||||||| ||  | | ||||    | |        |
Db        766 ATGAAGAAACTACAGATGAATAAATTAATAAATATAATTAATACATCACTGAATAAGTCA 825

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||| |   ||||  | |  | ||| |  ||||||||||||||   ||||||||||||||
Db        826 ATACAATCATCACCATTAAAAACTAATTTAGAAGAAAGTATGAAATATTCATTAAATGCT 885

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              || |||||| | ||| |||||||  | |||||| | ||  || |    || || |  ||
Db        886 GGTGGTAAAAGAATCAGACCAGTCATATTATTATTAACACTAAAAATGCTTAACAAAGAT 945

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              ||| |   |||  | || || ||  | ||| | ||||||||||||||| |||||  | |||
Db        946 TATCAACAAGGACTAAATAGTGCTTTAGCATTGGAAATGATTCATACTTATTCTTTAATT 1005

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
```

```
Db      1006 CATGATGATTTACCAGCAATGGATAATGACGATTACCGTAGAGGAAAATTAACAAATCAT 1065

Qy       301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Db      1066 AAAGTTTATGGTGAATGGAAAGCCATTCTTGCTGGTGATGCATTATTAACAAAAGCTTTT 1125

Qy       361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Db      1126 GAATTAGTTTCTAATGATACTACCATTGAAGATAGTGTGAAAGTAAGTATTATAAAAAGA 1185

Qy       421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Db      1186 CTTTCAAAAGCAAGTGGACATTTGGGAATGGTGGGTGGCCAAGCGCTTGATATGGAAAGT 1245

Qy       481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Db      1246 GAAGGGAAGTCAATTCGTTTAGAAACTTTAGAATCAATTCATGAAACTAAGACAGGCGCT 1305

Qy       541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600

Db      1306 TTACTAAATTTTTCAGTTATGGCTGCGGTAGACATTGCTCAAGTAGAACAAAATATTGCT 1365

Qy       601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660

Db      1366 AAGAATTTAGATGAATTTAGTCATCATTTAGGAATGATGTTTCAAATTAAAGATGATTTA 1425

Qy       661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Db      1426 CTGGATGTGTATGGTGATGAATCAAAACTTGGCAAAAAAGTAGGCAGTGATATAGTAAAT 1485

Qy       721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Db      1486 CATAAAAGTACTTATGTTTCTTTACTTGGAAAAGAAGGAGCAGAAGAAAAGTTAAACAAT 1545

Qy       781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840

Db      1546 CATCAATATCTTGCTATGAACTGCTTAAATCAAATTTCTGATCAATATGATACTTCTGAA 1605

Qy       841 TTATTAGAAATCGTTGATTTA 861

Db      1606 TTAAGTGATATTGTAGATTTA 1626


RESULT 7
AF269889/c
LOCUS       AF269889               4045 bp    DNA     linear   BCT 01-AUG-2000
DEFINITION  Staphylococcus epidermidis strain SR1 clone step.1028f08 genomic
            sequence.
ACCESSION   AF269889
VERSION     AF269889.1  GI:9623789
KEYWORDS    .
SOURCE      Staphylococcus epidermidis.
  ORGANISM  Staphylococcus epidermidis
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1  (bases 1 to 4045)
  AUTHORS   Kimmerly,W.J., Taylor,J.David., Nelsen,A.J., Godlevski,M.M.,
            Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
            Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,
            Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
            Furdon,P.J.
  TITLE     Transposon-mediated sequencing of the Staphylococcus epidermidis
            genome
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4045)
  AUTHORS   Taylor,J.David., Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M.,
            Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
            Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,
            Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
            Furdon,P.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAY-2000) Departments of Genomic Sciences and
            Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
            Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES             Location/Qualifiers
     source          1. .4045
                     /organism="Staphylococcus epidermidis"
                     /strain="SR1"
                     /db_xref="taxon:1282"
                     /clone="step.1028f08"
BASE COUNT     1249 a    691 c    536 g   1569 t
ORIGIN

  Query Match             52.2%;  Score 449.8;  DB 1;  Length 4045;
  Best Local Similarity   70.2%;  Pred. No. 6.2e-76;
  Matches  604;  Conservative    0;  Mismatches  257;  Indels    0;  Gaps    0;

Qy           1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Db        2313 ATGAAGAAACTACAGATGAATAAATTAATAAATATAATTAATACATCACTGAATAAGTCA 2254

Qy          61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120

Db        2253 ATACAATCATCACCATTAAAAACTAATTTAGAAGAAAGTATGAAATATTCATTAAATGCT 2194

Qy         121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Db        2193 GGTGGTAAAAGAATCAGACCAGTCATATTATTATTAACACTAAAAATGCTTAACAAAGAT 2134

Qy         181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Db        2133 TATCAACAAGGACTAAATAGTGCTTTAGCATTGGAAATGATTCATACTTATTCTTTAATT 2074

Qy         241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300

Db        2073 CATGATGATTTACCAGCAATGGATAATGACGATTACCGTAGAGGAAAATTAACAAATCAT 2014

Qy         301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Db        2013 AAAGTTTATGGTGAATGGAAAGCCATTCTTGCTGGTGATGCATTATTAACAAAAGCTTTT 1954

Qy         361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Db        1953 GAATTAGTTTCTAATGATACTACCATTGAAGATAGTGTGAAAGTAAGTATTATAAAAAGA 1894

Qy         421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Db        1893 CTTTCAAAAGCAAGTGGACATTTGGGAATGGTGGGTGGCCAAGCGCTTGATATGGAAAGT 1834

Qy         481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Db        1833 GAAGGGAAGTCAATTCGTTTAGAAACTTTAGAATCAATTCATGAAACTAAGACAGGCGCT 1774

Qy         541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600

Db        1773 TTACTAAATTTTTCAGTTATGGCTGCGGTAGACATTGCTCAAGTAGAACAAAATATTGCT 1714

Qy         601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660

Db        1713 AAGAATTTAGATGAATTTAGTCATCATTTAGGAATGATGTTTCAAATTAAAGATGATTTA 1654

Qy         661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Db        1653 CTGGATGTGTATGGTGATGAATCAAAACTTGGCAAAAAAGTAGGCAGTGATATAGTAAAT 1594

Qy         721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Db        1593 CATAAAAGTACTTATGTTTCTTTACTTGGAAAAGAAGGAGCAGAAGAAAAGTTAAACAAT 1534

Qy         781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840

Db        1533 CATCAATATCTTGCTATGAACTGCTTAAATCAAATTTCTGATCAATATGATACTTCTGAA 1474

Qy         841 TTATTAGAAATCGTTGATTTA 861

Db        1473 TTAAGTGATATTGTAGATTTA 1453
```

```
RESULT 8
AX145207/c
LOCUS       AX145207               4045 bp    DNA     linear   PAT 31-MAY-2001
DEFINITION  Sequence 3929 from Patent WO0134809.
ACCESSION   AX145207
VERSION     AX145207.1  GI:14283772
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4045)
  AUTHORS   Kimmerly,W.J.
  TITLE     Staphylococcus epidermidis nucleic acids and proteins
  JOURNAL   Patent: WO 0134809-A 3929 17-MAY-2001;
            GLAXO GROUP LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .4045
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic nucleic acid sequence"
BASE COUNT     1249 a    691 c    536 g   1569 t
ORIGIN

  Query Match           52.2%;  Score 449.8;  DB 6;  Length 4045;
  Best Local Similarity   70.2%;  Pred. No. 6.2e-76;
  Matches  604;  Conservative    0;  Mismatches  257;  Indels    0;  Gaps    0;

Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              |||| ||| |||| |||||||||||||||| ||  | | ||||    | |         |
Db       2313 ATGAAGAAACTACAGATGAATAAATTAATAAATATAATTAATACATCACTGAATAAGTCA 2254

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||| |   ||||  | |  | ||| |  ||||||||||||||   ||||||||||||||||
Db       2253 ATACAATCATCACCATTAAAAACTAATTTAGAAGAAAGTATGAAATATTCATTAAATGCT 2194

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              || |||||| | ||| |||||||  | |||||| | ||  || |    || || |  ||
Db       2193 GGTGGTAAAAGAATCAGACCAGTCATATTATTATTAACACTAAAAATGCTTAACAAAGAT 2134

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              ||| |   |||  | || || ||  | ||| | ||||||||||||| |||||  | |||
Db       2133 TATCAACAAGGACTAAATAGTGCTTTAGCATTGGAAATGATTCATACTTATTCTTTAATT 2074

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              ||||||||  |||||||  |||||||||||  |||||  ||  ||||||||||||||||||||
Db       2073 CATGATGATTTACCAGCAATGGATAATGACGATTACCGTAGAGGAAAATTAACAAATCAT 2014

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              |||||  |||||||| ||||  ||  ||  | ||  |||||||| |||||||| |||||  |||
Db       2013 AAAGTTTATGGTGAATGGAAAGCCATTCTTGCTGGTGATGCATTATTAACAAAAGCTTTT 1954

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              |||  |  |||| |  ||||  ||    |    |||   || ||| |||   || || || |
Db       1953 GAATTAGTTTCTAATGATACTACCATTGAAGATAGTGTGAAAGTAAGTATTATAAAAAGA 1894

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||  |||| |||||||||  ||| |  |||||||| || ||  |||   | |  ||||||  ||||
Db       1893 CTTTCAAAAGCAAGTGGACATTTGGGAATGGTGGGTGGCCAAGCGCTTGATATGGAAAGT 1834

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||  |  |||||  | |  |||||||| |||    || ||  |||| || ||||| ||
Db       1833 GAAGGGAAGTCAATTCGTTTAGAAACTTTAGAATCAATTCATGAAACTAAGACAGGCGCT 1774

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              |||  |||  |||| | ||||||   |||  | |||  || |||  | ||  ||   | |  | |
Db       1773 TTACTAAATTTTTCAGTTATGGCTGCGGTAGACATTGCTCAAGTAGAACAAAATATTGCT 1714

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
               |  |||||||    | |||| |||||||||| |||||||| || ||||||||||||||||
Db       1713 AAGAATTTAGATGAATTTAGTCATCATTTAGGAATGATGTTTCAAATTAAAGATGATTTA 1654

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              | ||    |||||||||||| ||||  | || |||||||| ||||| ||| | | ||||
Db       1653 CTGGATGTGTATGGTGATGAATCAAAACTTGGCAAAAAAGTAGGCAGTGATATAGTAAAT 1594

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||| || ||   |||| | || ||||| || |||||||| || || |   ||
Db       1593 CATAAAAGTACTTATGTTTCTTTACTTGGAAAAGAAGGAGCAGAAGAAAAGTTAAACAAT 1534

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              |||  | |    ||  || |    |||  ||||||  ||| ||||   ||||     |
Db       1533 CATCAATATCTTGCTATGAACTGCTTAAATCAAATTTCTGATCAATATGATACTTCTGAA 1474

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||   || || || ||||||
Db       1473 TTAAGTGATATTGTAGATTTA 1453


RESULT 9
AX141501
LOCUS       AX141501               783 bp    DNA     linear   PAT 31-MAY-2001
DEFINITION  Sequence 223 from Patent WO0134809.
ACCESSION   AX141501
VERSION     AX141501.1  GI:14281622
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 783)
  AUTHORS   Kimmerly,W.J.
  TITLE     Staphylococcus epidermidis nucleic acids and proteins
  JOURNAL   Patent: WO 0134809-A 223 17-MAY-2001;
            GLAXO GROUP LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .783
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic nucleic acid sequence"
BASE COUNT     297 a     93 c    147 g    246 t
ORIGIN

  Query Match           47.2%;  Score 406.8;  DB 6;  Length 783;
  Best Local Similarity   70.9%;  Pred. No. 1.2e-67;
  Matches  540;  Conservative    0;  Mismatches  222;  Indels    0;  Gaps    0;

Qy        100 ATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACT 159
              |||   |||||||||||||||| |||||| | ||| |||||||  | |||||| | ||
Db          1 ATGAAATATTCATTAAATGCTGGTGGTAAAAGAATCAGACCAGTCATATTATTATTAACA 60

Qy        160 TTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATG 219
              || |    || || |  || ||| |  |||  | || || ||  | ||| | ||||||
Db         61 CTAAAAATGCTTAACAAAGATTATCAACAAGGACTAAATAGTGCTTTAGCATTGGAAATG 120

Qy        220 ATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGA 279
              |||||||| |||||  | ||||||||||  |||||| |||||||||| ||||| ||
Db        121 ATTCATACTTATTCTTTAATTCATGATGATTTACCAGCAATGGATAATGACGATTACCGT 180

Qy        280 CGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGAT 339
              |||||||||||||||||||||||||| |||||||| ||||  || ||  | || ||||||
Db        181 AGAGGAAAATTAACAAATCATAAAGTTTATGGTGAATGGAAAGCCATTCTTGCTGGTGAT 240

Qy        340 GCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTA 399
              || |||||||| |||||  |||||| |  |||| | ||||  ||   |    |||   ||
Db        241 GCATTATTAACAAAAGCTTTTGAATTAGTTTCTAATGATACTACCATTGAAGATAGTGTG 300

Qy        400 AAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGT 459
              ||| |||   || || || | || |||| |||||||| ||| | |||||||| || ||
Db        301 AAAGTAAGTATTATAAAAAGACTTTCAAAAGCAAGTGGACATTTGGGAATGGTGGGTGGC 360

Qy        460 CAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATA 519
              ||| | | |||||| |||| |||||  |  |||||  | | |||||||| |||   ||
Db        361 CAAGCGCTTGATATGGAAAGTGAAGGGAAGTCAATTCGTTTAGAAACTTTAGAATCAATT 420

Qy        520 CACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCT 579
```

```
                ||  |||| || ||||| || ||| ||| |||| | ||||||  ||| | ||| || |||
Db        421 CATGAAACTAAGACAGGCGCTTTACTAAATTTTTCAGTTATGGCTGCGGTAGACATTGCT 480

Qy        580 AATGTCGATGATACAACTAAAGAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATG 639
                | || ||  | |  | |    |  |||||||    | |||| |||||||||| ||||||
Db        481 CAAGTAGAACAAAATATTGCTAAGAATTTAGATGAATTTAGTCATCATTTAGGAATGATG 540

Qy        640 TTCCAGATTAAAGATGATTTATTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAA 699
                || || |||||||||||||||| | ||    |||||||||||| ||||  | || ||||||
Db        541 TTTCAAATTAAAGATGATTTACTGGATGTGTATGGTGATGAATCAAAACTTGGCAAAAAA 600

Qy        700 GTGGGCAGCGATCTTGAAAATAATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGC 759
                || ||||| ||| | | |||| |||||||||| || ||   |||| | || ||||| ||
Db        601 GTAGGCAGTGATATAGTAAATCATAAAAGTACTTATGTTTCTTTACTTGGAAAAGAAGGA 660

Qy        760 GCAGAAGATAAATTGACTTATCATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGAT 819
                |||||||| || || |   |||||  | |    ||  || |     |||, ||||||  |
Db        661 GCAGAAGAAAAGTTAAACAATCATCAATATCTTGCTATGAACTGCTTAAATCAAATTTCT 720

Qy        820 GAACAATTCAATACAAAACACTTATTAGAAATCGTTGATTTA 861
                || ||||   ||||     | |||   || || || ||||||
Db        721 GATCAATATGATACTTCTGAATTAAGTGATATTGTAGATTTA 762


RESULT 10
AL591978
LOCUS       AL591978              250050 bp    DNA     linear  BCT 06-JUN-2002
DEFINITION  Listeria monocytogenes strain EGD, complete genome, segment 6/12.
ACCESSION   AL591978 AL591824
VERSION     AL591978.1  GI:16410540
KEYWORDS    .
SOURCE      Listeria monocytogenes.
  ORGANISM  Listeria monocytogenes
            Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
REFERENCE   1
  AUTHORS   Glaser,P., Frangeul,L., Buchrieser,C., Rusniok,C., Amend,A.,
            Baquero,F., Berche,P., Bloecker,H., Brandt,P., Chakraborty,T.,
            Charbit,A., Chetouani,F., Couve,E., de Daruvar,A., Dehoux,P.,
            Domann,E., Dominguez-Bernal,G., Duchaud,E., Durant,L.,
            Dussurget,O., Entian,K.D., Fsihi,H., Portillo,F.G., Garrido,P.,
            Gautier,L., Goebel,W., Gomez-Lopez,N., Hain,T., Hauf,J.,
            Jackson,D., Jones,L.M., Kaerst,U., Kreft,J., Kuhn,M., Kunst,F.,
            Kurapkat,G., Madueno,E., Maitournam,A., Vicente,J.M., Ng,E.,
            Nedjari,H., Nordsiek,G., Novella,S., de Pablos,B., Perez-Diaz,J.C.,
            Purcell,R., Remmel,B., Rose,M., Schlueter,T., Simoes,N.,
            Tierrez,A., Vazquez-Boland,J.A., Voss,H., Wehland,J. and Cossart,P.
  TITLE     Comparative genomics of Listeria species
  JOURNAL   Science 294 (5543), 849-852 (2001)
  MEDLINE   21537279
   PUBMED   11679669
REFERENCE   2  (bases 1 to 250050)
  AUTHORS   Glaser,P., Frangeul,L. and Rusniok,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUN-2001) Glaser P., Institut Pasteur, Genomique des
            Microorganismes Pathogenes, 25 rue du Docteur Roux, 75724 Paris
            Cedex 15, FRANCE
COMMENT     E-mail: pglaser@pasteur.fr
            Phone: +33 1 45 68 89 96, Fax: +33 (0)1 45 68 87 86.
FEATURES             Location/Qualifiers
     source          1. .250050
                     /organism="Listeria monocytogenes"
                     /strain="EGD-e"
                     /db_xref="taxon:1639"
     gene            210. .753
                     /gene="lmo1125"
     RBS             210. .215
                     /gene="lmo1125"
     CDS             220. .753
                     /gene="lmo1125"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC99203.1"
                     /db_xref="GI:16410541"
                     /db_xref="SPTREMBL:Q8Y7Z4"
                     /translation="MLIPSYLKSTAKELSINDFLNVEIVTINKEETFEILYCGTLKEI
                     EGDQLITGEAYEIPLKIIAKSTLSGKEILLYDGAYYGYDSMFCDEFEEEATQNRELKK
                     YPINNLSNIRLSIGIGIDYESEKEEYEFDENGHVVLIDDRHIPWEQVKTDGFDFLEIT
                     ATDENGDSLLILTEELA"
     gene            join(764. .769,778. .1230)
                     /gene="lmo1126"
     RBS             764. .769
                     /gene="lmo1126"
     CDS             778. .1230
                     /gene="lmo1126"
                     /note="similar to E. coli YjaB protein"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC99204.1"
                     /db_xref="GI:16410542"
                     /db_xref="SPTREMBL:Q8Y7Z3"
                     /translation="MILINKRAENQDYQMILAIWEKSVIATHDFLAAEDRQFYKEQIP
                     LFLDNVELFLWFAGEEIVGFSGTSERELDMLFLNPIATGNGYGSQILIWLIENKRINL
                     VDVNEQNENATRFYLKHGFVVSSRSDLDGFGNPYPILHLQLKERVNGI"
     gene            1208. .1213
                     /gene="lmo1127"
     RBS             1208. .1213
                     /gene="lmo1127"
     gene            1220. .1692
                     /gene="lmo1127"
     CDS             1220. .1675
                     /gene="lmo1127"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC99205.1"
                     /db_xref="GI:16410543"
                     /db_xref="SPTREMBL:Q8Y7Z2"
                     /translation="MEYEQLLQNELGELEIIQEELQNKEYRGMIVKTAKDGTTIRTRL
                     AKKTDKKAGYFVAFWEKDEFGKNCPFDENDCADFLCIVVIDGDLNGLFVFPRQCLIDK
                     GILTSEASIGKMAARFYPTWCQDLNKTAKKTQAWQQVYFKNYSIRKEDH"
     terminator      1660. .1692
                     /gene="lmo1127"
     gene            complement(1689. .2711)
                     /gene="lmo1128"
     CDS             complement(1689. .2711)
                     /gene="lmo1128"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC99206.1"
                     /db_xref="GI:16410544"
                     /db_xref="SPTREMBL:Q8Y7Z1"
                     /translation="MAKNGVDTIKRWKKITIIMSAIIVILVAILFIGNNYEMNETRVA
                     IPTTGGKLSAVVTTPKHGKPKGIIVFVHGDGAQEATQNGGYKPLMERFAKQGYISVSW
                     DKLGVGKSSGNWLNQSMDDRANEVNQVIEWMKVKYPDSTAKIGLWGASQAGWVVPKAM
                     NANNEIDFSILAAPAINWMRQGEYNTCWQAKDTGKTTKEVKQAQQNFLTDSALISKNK
                     TYESYKQNGGKEEMTPGRYNFIRKNMNQDATLDLSKVKTKIYLALAEKDKNVDSAETK
                     DVYTQQIKKENLETQTIPNVSHQMINPAIANSELLTNIVGLMVPKYFLVDNAYLDYCE
                     QVVSKQ"
     gene            complement(2666. .3280)
                     /gene="lmo1129"
     CDS             complement(2666. .3280)
                     /gene="lmo1129"
                     /note="similar to unknown proteins"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC99207.1"
                     /db_xref="GI:16410545"
                     /db_xref="SPTREMBL:Q8Y7Z0"
                     /translation="MTLSIKSIDSFEKDSIVEEVASLVFAGFKSKFTKTLFKETEAHD
                     ITYAFCHYIYTEKGENLLIAIQDELVCGCLFVTSKSDSYQKLYQTLRKFLSFSQCLKL
                     IFLLSLLSHKPVQNERYIDFITVSPNFRGQGIGKTLISHCLETFHDEKITLYVAKKNS
                     GAYKLYRNLGFQVIQKENSTLMGVLTGIKEWRKMEWIQLNVGKK"
     gene            complement(join(3277. .3996,4004. .4009))
                     /gene="lmo1130"
     CDS             complement(3277. .3996)
                     /gene="lmo1130"
                     /note="similar to transcription regulators"
```

```
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC99208.1"
                     /db_xref="GI:16410546"
                     /db_xref="SPTREMBL:Q8Y7Y9"
                     /translation="MLTLSTGEIAQLFQISKYKIRHYIDEGILVPKRNPENGYYYFEE
                     ADIYRLYQIILFRKIGFSIQEIKESLLGEKVTPMLEQAEQDLQQKIEELVEIQKTIQK
                     IIHSQKEMRLNEITFVAKDDRYFKKVPNQIIDHNTINYSQAAKHNLPDLEEPFYIFSK
                     QDTGVICLKSIKKRSDYTFPAGEYACKSFIAKDEATIETQIQLFLEELKYPTTSMILY
                     ENIYPSLAYPDAMVYTMEVPL"
gene                 complement(3288. .3293)
                     /gene="lmo1129"
RBS                  complement(3288. .3293)
                     /gene="lmo1129"
RBS                  complement(4004. .4009)
                     /gene="lmo1130"
gene                 join(4233. .4237,4245. .5960)
                     /gene="lmo1131"
RBS                  4233. .4237
                     /gene="lmo1131"
CDS                  4245. .5960
                     /gene="lmo1131"
                     /note="similar to ABC transporters, ATP-binding proteins"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC99209.1"
                     /db_xref="GI:16410547"
                     /db_xref="SPTREMBL:Q8Y7Y8"
                     /translation="MIDKRLFQLVEKKSLVLLILFRVLSLGLMIGLWLVFAQQLTHYL
                     EGQNVDWLWLIGTVLVVLVGKAILTKLAEKQIYQASAELRLSMRRAVMEKAFRLGNNE
                     GQLPASTLTQLAVDGIEQLEIYYSRFLPQLFYCLIASLMIFGSLVGFAWQPAIVLLIC
                     MPMIPIVIMAVMKIAKRILSGYWSDYTNLGTKFHENLSGLSILKAYDQDKYKQEEIVS
                     DAERFRKATMSLLSMQLNSITIMDIISYSGAALGIGMSLIMFTNGTISMTGMLMFLLL
                     SAEFFIPMRQLGSLFHVAMNGISACSKLFAYLELKEQVYGSVELAKPLEKIEVHHLTY
                     TYEGGKAKALQDVSVHFNKGSFSALVGKSGSGKSTFVRVLLNQLPGYQGEIVWNDVPL
                     ASLSGEAIRKQAVLVDNHGYLYANSIRENLLIGNPAANDSDLWNVLEQVSLADFVRKL
                     PEQLNENLEEDGSNLSGGQRQRLLLARALLRQAEVYVFDEITSGVDLESEKIILLVLQ
                     ELAKEKIVLFISHRLYNVLGADQVLVFDAGKLLEVASPERLQQESNYFKNYFLEEEAL
                     LKGGA"
gene                 5948. .5953
                     /gene="lmo1132"
RBS                  5948. .5953
                     /gene="lmo1132"
gene                 5962. .7608
                     /gene="lmo1132"
CDS                  5962. .7608
                     /gene="lmo1132"
                     /note="similar to ABC transporters, ATP-binding proteins"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC99210.1"
                     /db_xref="GI:16410548"
                     /db_xref="SPTREMBL:Q8Y7Y7"
                     /translation="MSEWTIIGWLLKFVKPLRGKMILAILLGIISNLSVIMISLIGTY
                     GIIAVILQQPLNPYKWLFVMVACGVVRGLARYLEQYLNHDIAFRLLAIIRERIFATLR
                     KLGPARLSGKKSGDLVAAITTDVEALEVFFAHTISPVFIALGTTIATVGFLASYDVGL
                     AIILLVGQIMVGVVLPMISYKRNKKIGTAYQQAFVGLNQMVMENIASLQDIFQFKLGE
                     ERLRKLETQGEKLNRQYKKRIRQGSELQILGEWVLIGTATIILVLGSLSQLPLETVLI
                     GTVLSLSSFGSVLALNALGTALLTTFASGKRLYALTEEKPVVLFNGQLELSDFESAEL
                     DKVSFSHDGKQPLLNELSLALPKGKWLGIGGESGSGKSTLVKLLMRYWDPDGQVKLNN
                     NELPKITESSLHRLEGVMEQRTFLFEDTLGNNIRLGKKDATLEEVKEAARKAAIDTWI
                     ETLPEGYETIIGGQSRNLSDGERQRIGLARLFLHDAPLLLLDEPTSNLDYINEQAILN
                     TLRSEIQDKTVLVISHRATTLDLAEEQLFIENGALKRAAK"
gene                 complement(join(7604. .7637,7645. .7953,7961. .7966))
                     /gene="lmo1133"
terminator           complement(7604. .7637)
                     /gene="lmo1133"
gene                 7615. .7637
                     /gene="lmo1132"
terminator           7615. .7637
                     /gene="lmo1132"
CDS                  complement(7645. .7953)
                     /gene="lmo1133"
                     /note="similar to B. subtilis YjcS protein"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC99211.1"
                     /db_xref="GI:16410549"
                     /db_xref="SPTREMBL:Q8Y7Y6"
                     /translation="MYLKENTNYLYCTAVIQTTGKVSYEQLVEHLEALRAKTVKEPGC
```

```
  Query Match             29.3%;  Score 252.6;  DB 1;  Length 250050;
  Best Local Similarity   59.3%;  Pred. No. 8.3e-39;
  Matches  429;  Conservative    0;  Mismatches  294;  Indels    0;  Gaps    0;

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||||| |   |   || ||  ||   ||| ||||    ||||| |||||  |  | ||
Db     226000 ATAAACGAGCGAAATATCGAACCTAGACTAAAAGAGTCCATGTTATATTCTATTCAAGCA 226059

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              || || ||||| || || ||| | ||  | ||    ||  |  |  |  ||||     |
Db     226060 GGCGGGAAACGGATTCGTCCAATGCTAGTTTTCGCTACGCTTCAAGCTTTAAAAGTAAAT 226119

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
                  |||||| | || | |||||  ||  ||||||||||||| |||||    ||||||
Db     226120 CCGCTTTTAGGTGTAAAAACCGCAACAGCGTTAGAAATGATTCACACATACAGCCTTATT 226179

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              |||||||| |||||||| |||||||| |||||||||||| ||||| || |  || ||||||
Db     226180 CATGATGATCTACCAGCAATGGATAACGATGATTATCGTCGAGGCAAGTGGACTAATCAT 226239

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||| | ||| ||    |||||||| || ||||| |||||||| ||||    || |||
Db     226240 AAAGTTTTTGGCGATGCAACTGCGATTTTGGCAGGAGATGCTTTACTAACGCTCGCTTTT 226299

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              || | |    || |||| ||| || |||  |   ||    |  | | | |
Db     226300 TCTATTTTAGCTGAAGACGATAATTTATCTTTTGAGACACGCATTGCTTTGATTAACCAA 226359

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              |    |    || |||   |  |||||||| || ||||||  |   |||  || ||
Db     226360 ATTAGTTTTAGTAGCGGTGCAGAAGGAATGGTTGGTGGTCAACTTGCAGACTTGGAAGCG 226419

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              |||  | |||  | |    || |||   ||   |   || ||   |  |||||| || | |
Db     226420 GAAAACAAACAAGTGACGCTAGAAGAGTTATCATCCATTCATGCACGAAAAACGGGTGAA 226479

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              ||||||| || ||| || |    |||||||  | || ||  | |  |||    |   |
Db     226480 TTATTAATTTATGCTGTAACCTCTGCAGCAAAAATTGCGGAAGCTGATCCAGAACAAACG 226539

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              |||  |||  || || |    |  || | || ||    || || ||||  || ||| |
Db     226540 AAACGCTTACGAATTTTTGCAGAGAATATTGGGATTGGATTTCAAATTAGCGACGATATT 226599

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              |||||      |||||||||| | ||  | ||||||||    ||  |||  |
Db     226600 TTAGATGTAATTGGTGATGAAACGAAAATGGGTAAAAAGACAGGGGCCGACGCTTTTCTG 226659

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              |||||||||| ||     | ||| |   |   ||||| ||| ||     ||| | | |
Db     226660 AATAAAAGTACCTATCCCGGATTACTCACGCTTGATGGGGCAAAAAGGGCATTAAATGAG 226719

Qy        781 CAT 783
              |||
Db     226720 CAT 226722
```

```
RESULT 11
AX416832
LOCUS       AX416832               5635 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 3823 from Patent WO0228891.
ACCESSION   AX416832
VERSION     AX416832.1  GI:21449341
KEYWORDS    .
SOURCE      Listeria monocytogenes ATCC 19115.
```

```
  ORGANISM  Listeria monocytogenes ATCC 19115
            Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
REFERENCE   1
  AUTHORS   Glaser,P. and Kunst,F.
  TITLE     Listeria innocua, genome and applications
  JOURNAL   Patent: WO 0228891-A 3823 11-APR-2002;
            Pasteur Institut (FR)
FEATURES             Location/Qualifiers
     source          1. .5635
                     /organism="Listeria monocytogenes ATCC 19115"
                     /db_xref="taxon:176281"
BASE COUNT     1896 a    927 c   1238 g   1574 t
ORIGIN
```

Query Match 27.9%; Score 239.8; DB 6; Length 5635;
Best Local Similarity 58.2%; Pred. No. 4.4e-36;
Matches 421; Conservative 0; Mismatches 302; Indels 0; Gaps 0;

```
Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||||   |   |   || ||  ||   ||| ||||    ||||| |||||  |  | ||
Db       2944 ATAAGCGAGCGAAATATCGAACCTAGACTAAAAGAGTCCATGTTATATTCTGTTCAAGCA 3003

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              || || ||||| || || ||  | ||  | ||    ||  |  |  | ||||     |
Db       3004 GGTGGAAAACGGATTCGTCCCATGCTAGTTTTTGCTACACTTCAAGCCTTAAAAGTAAAT 3063

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
                  |||||| |||| |  |||| ||  |||||||||||||||| |||   || ||
Db       3064 CCGCTTTTAGGTGTGAAAACAGCAACGGCGTTAGAAATGATTCATACGTATAGCCTGATC 3123

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              ||||||||  |||||||  ||||||||||||||| ||||| || || ||||  || ||||||
Db       3124 CATGATGATTTACCAGCAATGGATAATGATGACTATCGTCGCGGCAAATATACTAATCAT 3183

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              |||||  | |||  ||    |||||||| || ||||| || ||||||  ||||    || |||
Db       3184 AAAGTTTTTGGGGATGCAACTGCGATTTTGGCAGGAGACGCTTTACTAACGCTTGCTTTT 3243

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
                 ||  |  |    ||  ||  |  |||  ||  ||||  |   |||     |  |  | |
Db       3244 TCTATTTTAGCGGAAGACGAAAATTTATCTTTTGAAACACGCATAGCTTTGATTAATCAA 3303

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              |     |    ||||||    |  |||||||| |||||||| |   |||||||  |
Db       3304 ATTAGTTTTAGTAGTGGTGCAGAAGGAATGGTTGGCGGTCAACTTGCAGATATGGAGGCA 3363

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              |||  | |||  | |     | |||     |   |   || ||   |  |||||||| | |
Db       3364 GAAAACAAACAAGTGACATTAGAAGAACTATCATCCATCCATGCACGGAAAACAGGCGAA 3423

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              |||||||  ||||||  || |    ||| ||  | || ||  | |  ||      |   |
Db       3424 TTATTAATTTTTGCCGTAACTTCTGCTGCGAAAATTGCAGAAGCTGACCCGGAACAAACG 3483

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
                ||   |||   | || |    |  || | || ||    |||||  || |  ||  ||| |
Db       3484 AAAAGATTACGGATTTTTGCGGAAAATATTGGGATTGGATTCCAAATCAGCGACGATATT 3543

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              |||||      |||||||||| ||||  | ||||||||    ||   |||   |
Db       3544 TTAGATGTTATTGGTGATGAAACAAAAATGGGTAAAAAGACAGGGGTCGACGCTTTTCTG 3603

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              || |||||||| ||      | ||| |   |   |||||  ||  ||     ||| | | |
Db       3604 AACAAAAGTACCTATCCCGGATTACTCACGCTTGATGGTGCCAAAAGGGCATTAAATGAG 3663

Qy        781 CAT 783
              |||
Db       3664 CAT 3666
```

RESULT 12
AB003187

```
LOCUS       AB003187               1080 bp    DNA     linear   BCT 25-MAR-1998
DEFINITION  Micrococcus luteus DNA for farnesyl diphosphate synthase, complete
            cds.
ACCESSION   AB003187
VERSION     AB003187.1  GI:2982676
KEYWORDS    farnesyl diphosphate synthase.
SOURCE      Micrococcus luteus (strain:B-P 26) DNA.
  ORGANISM  Micrococcus luteus
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Micrococcineae; Micrococcaceae; Micrococcus.
REFERENCE   1  (sites)
  AUTHORS   Shimizu,N., Koyama,T. and Ogura,K.
  TITLE     Molecular cloning, expression, and characterization of the genes
            encoding the two essential protein components of Micrococcus luteus
            B-P 26 hexaprenyl diphosphate synthase
  JOURNAL   J. Bacteriol. 180 (6), 1578-1581 (1998)
  MEDLINE   98175686
REFERENCE   2  (bases 1 to 1080)
  AUTHORS   koyama,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-APR-1997) Tanetoshi koyama, Tohoku University,
            Institute for Chemical Reaction Science; Katahira 2-1-1, Aoba-ku,
            Sendai, Miyagi 980-77, Japan (E-mail:koyama@icrs.tohoku.ac.jp,
            Tel:+81-22-217-5621, Fax:+81-22-217-5620)
FEATURES             Location/Qualifiers
     source          1. .1080
                     /organism="Micrococcus luteus"
                     /strain="B-P 26"
                     /db_xref="taxon:1270"
     gene            105. .980
                     /gene="fps"
     CDS             105. .980
                     /gene="fps"
                     /codon_start=1
                     /transl_table=11
                     /product="farnesyl diphosphate synthase"
                     /protein_id="BAA25265.1"
                     /db_xref="GI:2982677"
                     /translation="MLQEKLTMNRDFLNLINESLLNKYHPAQSRLHEAINYSLSAGGK
                     RIRPLLVLTTLDSLGGNAHDGLPPGIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHK
                     RFDEATAILAGDALLTDAFQCILNTQLNAEIKLSLINLLSTASGSNGMVYGQMLDMQG
                     EHKTLTLNELERIHIHKTGELIRAAIVSAGIIMNFNDAQIEQLNIIGKNVGLMFQIKD
                     DILDVEGSFENIGKTVGSDLNNDKSTYVSLLGLEASKQLLNDKLTETYDALKTLQPIN
                     DNLKTLITYIVERNK"
BASE COUNT      377 a    161 c    198 g    344 t
ORIGIN
```

Query Match 27.1%; Score 233.4; DB 1; Length 1080;
Best Local Similarity 57.4%; Pred. No. 9.6e-35;
Matches 461; Conservative 0; Mismatches 336; Indels 6; Gaps 2;

```
Qy          6 GAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAA 65
              ||| ||  | ||||||       |   | |  | | ||||||||| || | |    | | |
Db        116 GAAACTGACAATGAATCGTGACTTTCTTAATCTGATTAATGAATCATTGTTAAATAAATA 175

Qy         66 TAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGG 125
              | |  | |    | |  |    ||| | |||   ||    |||||||||| ||| || ||
Db        176 TCATCCTG---CGCAATCAAGACTACATGAAGCAATCAACTATTCATTAAGTGCAGGTGG 232

Qy        126 TAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGA 185
              || |||||  ||||| | ||  |  |    |||||||||     ||  |  | |      |
Db        233 CAAGCGCATTAGACCATTACTTGTTCTGACTACTTTAGATAGTTTAGGTGGCAATGCACA 292

Qy        186 GTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGA 245
                 |||  |     |  ||||| || ||||||||||||||| |||||  | |||||  ||
Db        293 TGACGGTTTACCATTTGGCATTGCGCTTGAAATGATTCATACGTATTCTTTAATTCACGA 352

Qy        246 TGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGT 305
              ||||  | || || ||||||||||||||| ||||| || || |||  | || ||||||||
Db        353 TGACTTGCCGGCAATGGATAATGATGACTATCGTCGCGGTAAACTCACGAATCATAAGCG 412

Qy        306 ATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACT 365
               | || |||    || || ||| | || || ||||| ||  | ||| | || ||| ||
```

```
Db        413 TTTTGATGAAGCAACAGCTATACTCGCTGGAGATGCATTGCTCACTGATGCTTTTCAATG 472

Qy        366 TATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTC 425
              |||| ||   ||      |||| |  ||| | ||| ||  | | |  |      |
Db        473 CATTTTAA---ATACGCAGTTAAACGCAGAAATTAAATTATCATTGATTAATTTATTAAG 529

Qy        426 AATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGG 485
              |  ||   |||   | ||| |||||   ||| |||||| | |||||||||| | |||
Db        530 TACTGCTTCTGGATCTAATGGCATGGTTTACGGCCAAATGCTCGATATGCAAGGTGAACA 589

Qy        486 CCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATT 545
              || || |     |  |     |||||   || |||| |   |||||| || | |||  |
Db        590 TAAAACATTGACATTAAATGAACTGGAACGTATTCACATACATAAAACCGGTGAATTGAT 649

Qy        546 AACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACA 605
              |   ||  || | |||||||     ||    ||| |  |||||  ||   |  || ||
Db        650 TCGTGCAGCAATTGTAAGTGCAGGTATCATAATGAATTTTAATGATGCACAAATTGAGCA 709

Qy        606 TTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGA 665
              |  | | |  | || |  |  |||||  | ||||||||||| ||||||||| |  | ||
Db        710 ACTCAATATTATTGGTAAGAACGTAGGTTTAATGTTCCAGATAAAAGATGATATTCTTGA 769

Qy        666 CTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAA 725
              | ||       | |||  | || |||| ||| ||||| ||| |   | ||| ||||
Db        770 TGTTGAAGGGTCGTTCGAAAATATCGGAAAAACAGTTGGCAGTGATTTAAATAATGATAA 829

Qy        726 AAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTATCATAG 785
              |||||| || ||     | |||||   || |  || |  | | | | || |
Db        830 AAGTACATATGTATCACTCTTAGGACTTGAAGCTTCAAAGCAATTACTTAATGATAAGCT 889

Qy        786 AGACGCAGCAGTGGATGAACTAA 808
              |   | | ||   |||| || |
Db        890 AACTGAAACATATGATGCGCTTA 912


RESULT 13
AX417035/c
LOCUS       AX417035              195269 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 4026 from Patent WO0228891.
ACCESSION   AX417035
VERSION     AX417035.1  GI:21449645
KEYWORDS    .
SOURCE      Listeria innocua.
  ORGANISM  Listeria innocua
            Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
REFERENCE   1
  AUTHORS   Glaser,P. and Kunst,F.
  TITLE     Listeria innocua, genome and applications
  JOURNAL   Patent: WO 0228891-A 4026 11-APR-2002;
            Pasteur Institut (FR)
FEATURES             Location/Qualifiers
     source          1. .195269
                     /organism="Listeria innocua"
                     /db_xref="taxon:1642"
                     /note="second part of seq 8-first part = seq 8 (000001 -
                     349980)-second part = seq 4026 (300001 - 495269)"
BASE COUNT    57412 a  39214 c  31781 g  66862 t
ORIGIN

  Query Match           26.9%;  Score 231.2;  DB 6;  Length 195269;
  Best Local Similarity   56.4%;  Pred. No. 1e-34;
  Matches  431;  Conservative    0;  Mismatches  333;  Indels    0;  Gaps    0;

Qy         20 ATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTAATGG 79
              || | |  | | | | | |  || |     | |     | |||||||        || |
Db      66205 ATCATTATAAAAAAGTACTTGATGAGTCGCTTTTTAAAGAAATAAATATGCGCAATATCG 66146

Qy         80 ATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCGAC 139
              |  || || |  |||||   ||||| |||||| |  | ||||| || ||||| || || |
Db      66145 AACCTAAGTTGAAAGAATCAATGTTATATTCAGTGCAAGCTGGCGGAAAACGAATTCGTC 66086

Qy        140 CAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGA 199
              || | ||  | ||    ||  |  |  |  ||||||  |||     |  ||| | || |
Db      66085 CAATGCTTGTTTTTGCAACGCTTCAAGCCTTAAATATTGAGCCAATGCGCGGTTTAAAAA 66026

Qy        200 GCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAGCGA 259
              || |  ||||| ||||||||||||||| ||      | ||||| |||||  ||||||| |
Db      66025 CAGCTACGGCACTTGAAATGATTCATACGTACAGCTTAATTCACGATGATTTACCAGCAA 65966

Qy        260 TGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGA 319
              ||||| ||||||| |||||||| || ||||  || || ||||||  |||| |||||    |
Db      65965 TGGATGATGATGACTATCGACGCGGGAAATGGACGAACCATAAAATATACGGTGATGCAA 65906

Qy        320 CTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATG 379
              | || || |||||||| |||||||| |||||     || |||     || |  |    ||||
Db      65905 CAGCAATTTTAGCAGGAGATGCTTTGTTAACACTCGCTTTTTCTATTTTAGCTGAAGATG 65846

Qy        380 ATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTC 439
              | |   || ||  |||      ||      |  |  | |  |      |||||
Db      65845 AAAATCTATCTTTCGAAACGCGTATTGCTTTAATTAATCAAATTAGTTATAGCAGTGGAG 65786

Qy        440 ATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATC 499
              |  || ||||| || || |||       ||||||  ||   |||    ||| | |
Db      65785 CAGAAGGCATGGTAGGTGGGCAACAAGCCGATATGGAAGCTGAAAATAAACAAGTCACAT 65726

Qy        500 TTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTA 559
              | |||    | | |   || |||       |||||  || | | |||||| ||| || || |
Db      65725 TAGAAGAACTAGCATCAATCCACGCTCGCAAAACTGGTGAACTATTAATTTTCGCAGTAA 65666

Qy        560 TGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAGAAAGTTATA 619
              || ||| | ||||||  | |     |     |   | | |||  |||  || || |
Db      65665 CCTCAGCCGCAAAAATCGCTGAAGCAACTCCAGAACAAACAAAACGATTACGAATTTTTG 65606

Qy        620 GTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCTATGGTGATG 679
              |  || | || ||    || || ||||  || ||| | |||||      ||| ||||
Db      65605 CAGAAAATATCGGCATTGGTTTTCAAATTAGCGACGATATTTTAGATGTAATTGGCGATG 65546

Qy        680 AAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTACGTACGTGA 739
              || ||||  | ||||||||    ||   |||   |       |||||||||| ||
Db      65545 AAACAAAAATGGGTAAAAAGACAGGGGTCGACGCTTTTCTGAATAAAAGTACCTATCCCG 65486

Qy        740 GTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTATCAT 783
              | ||| |       || || ||  ||     ||| | | | |||
Db      65485 GATTACTCACACTCGAAGGTGCTAAACGGGCATTAAATGAGCAT 65442


RESULT 14
AL596168
LOCUS       AL596168              333050 bp    DNA     linear   BCT 06-JUN-2002
DEFINITION  Listeria innocua Clip11262 complete genome, segment 6/12.
ACCESSION   AL596168 AL592022
VERSION     AL596168.1  GI:16413677
KEYWORDS    .
SOURCE      Listeria innocua.
  ORGANISM  Listeria innocua
            Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
REFERENCE   1
  AUTHORS   Glaser,P., Frangeul,L., Buchrieser,C., Rusniok,C., Amend,A.,
            Baquero,F., Berche,P., Bloecker,H., Brandt,P., Chakraborty,T.,
            Charbit,A., Chetouani,F., Couve,E., de Daruvar,A., Dehoux,P.,
            Domann,E., Dominguez-Bernal,G., Duchaud,E., Durant,L.,
            Dussurget,O., Entian,K.D., Fsihi,H., Portillo,F.G., Garrido,P.,
            Gautier,L., Goebel,W., Gomez-Lopez,N., Hain,T., Hauf,J.,
            Jackson,D., Jones,L.M., Kaerst,U., Kreft,J., Kuhn,M., Kunst,F.,
            Kurapkat,G., Madueno,E., Maitournam,A., Vicente,J.M., Ng,E.,
            Nedjari,H., Nordsiek,G., Novella,S., de Pablos,B., Perez-Diaz,J.C.,
            Purcell,R., Remmel,B., Rose,M., Schlueter,T., Simoes,N.,
            Tierrez,A., Vazquez-Boland,J.A., Voss,H., Wehland,J. and Cossart,P.
  TITLE     Comparative genomics of Listeria species
  JOURNAL   Science 294 (5543), 849-852 (2001)
  MEDLINE   21537279
   PUBMED   11679669
REFERENCE   2  (bases 1 to 333050)
  AUTHORS   Glaser,P., Frangeul,L. and Rusniok,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-2001) Glaser P., Institut Pasteur, Genomique des
```

```
          Microorganismes Pathogenes, 25 rue du Docteur Roux, 75724 Paris
          Cedex 15, FRANCE
COMMENT   E-mail: pglaser@pasteur.fr
          Phone: +33 (0)1 45 68 89 96, Fax: +33 (0)1 45 68 87 86.
FEATURES             Location/Qualifiers
     source          1. .333050
                     /organism="Listeria innocua"
                     /strain="Clip11262"
                     /db_xref="taxon:1642"
     gene            915. .4679
                     /gene="lin1204"
     RBS             915. .920
                     /gene="lin1204"
     CDS             927. .4640
                     /gene="lin1204"
                     /note="similar to internalin proteins (LPXTG motif)"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC96435.1"
                     /db_xref="GI:16413678"
                     /db_xref="SPTREMBL:Q92CG8"
                     /translation="MKRKISSIIVVGIMLFQSLTMYPFITEAKENEQKEEINKPSKIT
                     KGLTNSLKYTKTILETGDTYDSVFPDSALAKVVAKEATGSENTTQLVTQADLNKIKSL
                     NGYNKGISVLTGIDLLVNVTSISLNNNQVTDISPIDQLPNLVSLSVKNNQISSLILNA
                     QNQLPKLTTIDIENNPDLNTIDIQDQPQLVDVKTSGYTGLRKLTTVIAKNNPELVNLG
                     QYTIRNVYFSQVASLTKVELVNLPKVRKVNLERNSINELKVTDLAIEDLPLEENELTD
                     TVFDNIQNLPNLKTLDLSKNQLEEVVLDKTDVENLPNLMTLDIQQNLAIKLINVQDQP
                     QLVDVKTSDYKELSALTTVIAKNNPELVNLGQYTIRNVYFSQVASLTKVELVNLPKVR
                     KVNLERNSINELKVTDLAIEDLPLEENELTDTVFDNIQNLPNLKTLDLSKNQLEEVVL
                     DKTDVENLPNLMTLDIQQNLAIKLINVQDQPQLVDVKTSDYKELSALTTVIAKNNPEL
                     VNLGYPIMQNVYFSQVASLTKVELANLPKVRAVRLERNSINQIELNNLVSVKDVNLNT
                     NKITNDSIEKFRGMPILATLNLNKNQITNINMLDGFPQMTTLNIDLNSISVLPSNLKT
                     KMPKLSRISALNQTVTLGKAIVVDDSDLIINNEISNFEKLTEPSPISNFGTYANEKIT
                     WSSERIKNLTEVSFQFSELINVTGIDGTFSGKVTQPFKRSTTPVINADSEIHYPQGTE
                     KTEAEFLKDVQAQTTDDLAIKSDFETMVNLKKVGKYTVTLNVENLDGIKATPKEVIVY
                     IDAVQGANITVKYEDESGNKLAENSILTGNVGEEYSSSAKEILGYTLTEMPSNAQGEF
                     SLEEQTVTYIYSKNPVPAKDITVQYTDEDGMELAPTETLSGNVDENYVTTAKTFTGYE
                     LIETPSNAEGKFSENAQTVTYVYRAIKAEPILAKEVTVNYQDELGAKISETEVLTGEI
                     GETYTTVAKTIDGYTLIKSPTNANGIFNENPQTVTYVYQLQNSPIAANITVKHLDENN
                     NELAPSEVLSGNVDEAYTTKPKEIKDYSLVKVPTNASGKFTTEAQTVIYHYKKNSIPT
                     SSYITVKYVDETGKELAISKVLNGNINDSYATTAKEIKGYTLVQKPANATGKFTDQNQ
                     TIKYVYRANKAEVNPNPNSDVLGSNSNGDIQPTNDIKPTNDVQPTNEAPQTLPKTGSQ
                     PANLIFGLGILLVLLSTQWLHRNKRKKAKSHNC"
     terminator      4659. .4679
                     /gene="lin1204"
     gene            join(4883. .6241,6249. .6276)
                     /gene="lin1205"
     CDS             4883. .6241
                     /gene="lin1205"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC96436.1"
                     /db_xref="GI:16413679"
                     /db_xref="SPTREMBL:Q92CG7"
                     /translation="MKKMLPESKVESIRKEGFLNRAVEAYRFFYPTVSNVSNFKALHD
                     LGITENHDFIIQLTTPDLNVLTQNSDTPYCLGTGNTENGPVVIELPLGAIVGVADDIN
                     FKFITNMGLTGDEEGKGAKYLYLPPNYDGDIPEGYIVRKPSSFRFLICLRAMVHQASD
                     YEKAFELLKTVKFYPLKEAKNNPTSTFHDFSHRKAISTPYYVEGKFDYWEVIKWALDN
                     DETDPEYYQMYGLLKAIGLAPNKEFNPDSDTKALLIEAAEKADKMMFVNSFNTDDPAA
                     IVWPGKNWEWAVYGENNDFYEKTYLNLPVRERWFYQATLETHMMFMHKVGFGSVYMLG
                     VKDKEGNYLDGGKSYTLKVPTPVPTSIFWSVTVYEMDTRSEIVTEQFMPALNSIKDTF
                     EADADGNTTLYFGPNPPEDESLPWIQTVPDANWFTYFRIYGPTEPAFDNSWQLYDFEE
                     VK"
     gene            complement(join(6249. .6276,6283. .6876,6885. .6890))
                     /gene="lin1206"
     terminator      6249. .6276
                     /gene="lin1205"
     terminator      complement(6249. .6276)
                     /gene="lin1206"
     CDS             complement(6283. .6876)
                     /gene="lin1206"
                     /note="similar to B. subtilis YdeI protein"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC96437.1"
                     /db_xref="GI:16413680"
                     /db_xref="SPTREMBL:Q92CG6"
                     /translation="MAKTELNPKVDSFLNKPSTWQAEFKALREMAITFELEEEFKWGK
                     PCYAINGSNVFLIHGFKNYCALLFMKGALLRDPDNILVQQTENVQAARQIRFTNLQEI
                     LDQKEILKKYIQNAIEVEKAGLEVALKPRAETPIPEELLVKFEELPALKTAFEALTPG
                     RQKAYLLYFAAPKQSKTRVSRIEKYEATILDGLGLND"
     RBS             complement(6885. .6890)
                     /gene="lin1206"
     gene            complement(6994. .7414)
                     /gene="lin1207"
     CDS             complement(6994. .7401)
                     /gene="lin1207"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC96438.1"
                     /db_xref="GI:16413681"
                     /db_xref="SPTREMBL:Q92CG5"
                     /translation="MTLNVYLNFRTQSRDAIAFYEEIFGTKCTDLMTYGEIESAEEPI
                     EDSLKDLVMNASLVLDGVKVMFSDVPKSMPLTFGDNITLVIDTTDEIKLTKQFHQLAE
                     GGNIIMPLGKTFWSEKYGQVTDKFGIGWQLNLS"
     RBS             complement(7409. .7414)
                     /gene="lin1207"
     gene            7553. .8165
                     /gene="lin1208"
     RBS             7553. .7558
                     /gene="lin1208"
     CDS             7566. .8165
                     /gene="lin1208"
                     /note="weakly similar to phosphoglycerate mutase 1"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC96439.1"
                     /db_xref="GI:16413682"
                     /db_xref="SPTREMBL:Q92CG4"
                     /translation="MKKTLYLMRHGQTLFNQRKKIQGFCDAPLTDLGIKQAKIAGSYF
                     KENNITFDQVYSSTSERACDTLELITDKSYQRLKGLKEWNFGAFEGESEDLNPPLPYG
                     DFFAEYGGEREVDFRDRLVTTMERIMSQENHDTVLAVSHGAACAQFARYWEKTSKIGK
                     VTGLKNCCILKFEYENGRFTLVNFINHDFENGTHMESVK"
     gene            join(8347. .8352,8361. .8723)
                     /gene="lin1209"
     RBS             8347. .8352
                     /gene="lin1209"
     CDS             8361. .8723
                     /gene="lin1209"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC96440.1"
                     /db_xref="GI:16413683"
                     /db_xref="SPTREMBL:Q92CG3"
                     /translation="MSKYDAEIVNAQQSITGQQEKIRRLKELITKIERKGEQISSIPI
                     RQIDLPFDFWQGKSDSVMREVLQRRLQFWNNQNAVVNDLVQELRAAMTKMQNQISDYQ
                     ADISYYTNLKQMEEEVNV"
     gene            8701. .8706
                     /gene="lin1210"
     RBS             8701. .8706
                     /gene="lin1210"
     gene            8716. .8991
                     /gene="lin1210"
     CDS             8716. .8991
                     /gene="lin1210"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC96441.1"
                     /db_xref="GI:16413684"
                     /db_xref="SPTREMBL:Q92CG2"
                     /translation="MSDIQVNESQIIKSMNQFAGKQKVLYKKSSSTSHNNKLATYHKI
                     QTILTKMDTQMNQLQQLGERTTQDIEKNCAEFVNLDKNLLTDIEVTG"
     gene            8992. .10532
                     /gene="lin1211"
     RBS             8992. .8997
                     /gene="lin1211"
```

```
     CDS             8999. .10459
                     /gene="lin1211"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC96442.1"
                     /db_xref="GI:16413685"
                     /db_xref="SPTREMBL:Q92CG1"
                     /translation="MKIDVRELNEIVSENIVMLKKEYENLNIAMNATTNFTVETQEFF
                     KGKTADASRAYLQEAYKPVQQKALEVNKLMTKILSEYIADAEAQFGANGMIDIQSIEM
                     EYKRSFNQVSDDEMREYRDLNNLIQEANEFTYFQSSNLGFFEELHHDALTEITNIRMK
                     LEDFETKWNVELSKAEKLQKELDWMLTQVNDNKIIPTSYQAGTLPFGKPFKSRDEFDT
                     VQEYYQYVVEWAVANGKAKEIIVDGQRYYEFTDTFYYENYNKEVNGTLVMTEVNGEMI
                     LFVQRKDIESAVSVPFGVFYPDPFQRLANGGETDGTIKLYHWKSKDIPEGLSVPNEEL
                     KAMADFFADTQNSYDAGMAILKYLAHNKNPEKLSFTDTELDKVATNWSGGKNRFDFAG
                     EVEAHADMLTDSILTELAAKPMDKIDRPGNTDEEKEQGYYHAASIADMQKGEDGMMNK
                     SFNSIWGSGQYSAYLQNLMHGQENTPKAKKRSFPRQ"
     terminator      10513. .10532
                     /gene="lin1211"
     gene            join(10896. .11450,11456. .11489)
                     /gene="lin1212"
     CDS             10896. .11450
                     /gene="lin1212"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC96443.1"
                     /db_xref="GI:16413686"

  Query Match             26.9%;  Score 231.2;  DB 1;  Length 333050;
  Best Local Similarity   56.4%;  Pred. No. 9.1e-35;
  Matches  431;  Conservative    0;  Mismatches  333;  Indels    0;  Gaps    0;

Qy         20 ATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTAATGG 79
Db     174746 ATCATTATAAAAAAGTACTTGATGAGTCGCTTTTTAAAGAAATAAATATGCGCAATATCG 174805

Qy         80 ATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCGAC 139
Db     174806 AACCTAAGTTGAAAGAATCAATGTTATATTCAGTGCAAGCTGGCGGAAAACGAATTCGTC 174865

Qy        140 CAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGA 199
Db     174866 CAATGCTTGTTTTTGCAACGCTTCAAGCCTTAAATATTGAGCCAATGCGCGGTTTAAAAA 174925

Qy        200 GCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAGCGA 259
Db     174926 CAGCTACGGCACTTGAAATGATTCATACGTACAGCTTAATTCACGATGATTTACCAGCAA 174985

           260 TGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGA 319
Db     174986 TGGATGATGATGACTATCGACGCGGGAAATGGACGAACCATAAAATATACGGTGATGCAA 175045

Qy        320 CTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATG 379
Db     175046 CAGCAATTTTAGCAGGAGATGCTTTGTTAACACTCGCTTTTTCTATTTTAGCTGAAGATG 175105

Qy        380 ATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTC 439
Db     175106 AAAATCTATCTTTCGAAACGCGTATTGCTTTAATTAATCAAATTAGTTATAGCAGTGGAG 175165

Qy        440 ATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATC 499
Db     175166 CAGAAGGCATGGTAGGTGGGCAACAAGCCGATATGGAAGCTGAAAATAAACAAGTCACAT 175225

Qy        500 TTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTA 559
Db     175226 TAGAAGAACTAGCATCAATCCACGCTCGCAAAACTGGTGAACTATTAATTTTCGCAGTAA 175285

Qy        560 TGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAGAAAGTTATA 619
Db     175286 CCTCAGCCGCAAAAATCGCTGAAGCAACTCCAGAACAAACAAAACGATTACGAATTTTTG 175345

Qy        620 GTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCTATGGTGATG 679
Db     175346 CAGAAAATATCGGCATTGGTTTTCAAATTAGCGACGATATTTTAGATGTAATTGGCGATG 175405

Qy        680 AAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTACGTACGTGA 739
Db     175406 AAACAAAAATGGGTAAAAAGACAGGGGTCGACGCTTTTCTGAATAAAAGTACCTATCCCG 175465

Qy        740 GTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTATCAT 783
Db     175466 GATTACTCACACTCGAAGGTGCTAAACGGGCATTAAATGAGCAT 175509


RESULT 15
AX417044
LOCUS       AX417044               349980 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 4035 from Patent WO0228891.
ACCESSION   AX417044
VERSION     AX417044.1  GI:21449654
KEYWORDS    .
SOURCE      Listeria innocua.
  ORGANISM  Listeria innocua
            Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
REFERENCE   1
  AUTHORS   Glaser,P. and Kunst,F.
  TITLE     Listeria innocua, genome and applications
  JOURNAL   Patent: WO 0228891-A 4035 11-APR-2002;
            Pasteur Institut (FR)
FEATURES             Location/Qualifiers
     source          1. .349980
                     /organism="Listeria innocua"
                     /db_xref="taxon:1642"
                     /note="seq 2058, original length: 3.011.208 replaced
                     by-seq 2058: 0.000.001 to 0.349.980-seq 4032: 0.300.001 to
                     0.649.980-seq 4033: 0.600.001 to 0.949.980-seq 4034:
                     0.900.001 to 1.249.980-seq 4035: 1.200.001 to
                     1.549.980-seq 4036: 1.500.001 to 1.849.980-seq 4037:
                     1.800.001 to 2.149.980-seq 4038: 2.100.001 to
                     2.449.980-seq 4039: 2.400.001 to 2.749.980-seq 4040:
                     2.700.001 to 3.049.980-seq 4041: 3.000.001 to 3.011.208"
BASE COUNT   116021 a  60232 c  66706 g 107021 t
ORIGIN

  Query Match             26.9%;  Score 231.2;  DB 6;  Length 349980;
  Best Local Similarity   56.4%;  Pred. No. 9e-35;
  Matches  431;  Conservative    0;  Mismatches  333;  Indels    0;  Gaps    0;

Qy         20 ATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTAATGG 79
Db     192746 ATCATTATAAAAAAGTACTTGATGAGTCGCTTTTTAAAGAAATAAATATGCGCAATATCG 192805

Qy         80 ATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCGAC 139
Db     192806 AACCTAAGTTGAAAGAATCAATGTTATATTCAGTGCAAGCTGGCGGAAAACGAATTCGTC 192865

Qy        140 CAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGA 199
Db     192866 CAATGCTTGTTTTTGCAACGCTTCAAGCCTTAAATATTGAGCCAATGCGCGGTTTAAAAA 192925

Qy        200 GCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAGCGA 259
Db     192926 CAGCTACGGCACTTGAAATGATTCATACGTACAGCTTAATTCACGATGATTTACCAGCAA 192985

Qy        260 TGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGA 319
Db     192986 TGGATGATGATGACTATCGACGCGGGAAATGGACGAACCATAAAATATACGGTGATGCAA 193045

Qy        320 CTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATG 379
Db     193046 CAGCAATTTTAGCAGGAGATGCTTTGTTAACACTCGCTTTTTCTATTTTAGCTGAAGATG 193105

Qy        380 ATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTC 439
Db     193106 AAAATCTATCTTTCGAAACGCGTATTGCTTTAATTAATCAAATTAGTTATAGCAGTGGAG 193165
```

```
Qy        440 ATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATC 499
              |  || ||||| || || |||      ||||||  ||   |||    ||| | |
Db     193166 CAGAAGGCATGGTAGGTGGGCAACAAGCCGATATGGAAGCTGAAAATAAACAAGTCACAT 193225

Qy        500 TTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTA 559
              | |||    | | |   || |||      |||||  || | | ||||||  ||| || || |
Db     193226 TAGAAGAACTAGCATCAATCCACGCTCGCAAAACTGGTGAACTATTAATTTTCGCAGTAA 193285

Qy        560 TGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAGAAAGTTATA 619
                || ||| | |||||| | |    |     |   | | |||  |||  || || |
Db     193286 CCTCAGCCGCAAAAATCGCTGAAGCAACTCCAGAACAAACAAAACGATTACGAATTTTTG 193345

Qy        620 GTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCTATGGTGATG 679
              |  || | || ||    || || ||||  || ||| | |||||      ||| ||||
Db     193346 CAGAAAATATCGGCATTGGTTTTCAAATTAGCGACGATATTTTAGATGTAATTGGCGATG 193405

Qy        680 AAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTACGTACGTGA 739
              || ||||  | ||||||||    ||   |||   |      ||||||||||| ||
Db     193406 AAACAAAAATGGGTAAAAAGACAGGGGTCGACGCTTTTCTGAATAAAAGTACCTATCCCG 193465

Qy        740 GTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTATCAT 783
              | ||| |       || || ||  ||     ||| | | | |||
Db     193466 GATTACTCACACTCGAAGGTGCTAAACGGGCATTAAATGAGCAT 193509
```

Search completed: May 30, 2003, 16:56:52
Job time : 2487 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 30, 2003, 13:05:08 ; Search time 280 Seconds
(without alignments)
6924.895 Million cell updates/sec

Title: US-09-925-637-63
Perfect score: 861
Sequence: 1 atgacgaatctaccgatgaa.....tattagaaatcgttgattta 861

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		861	100.0	861	22	AAS00821	S. aureus HGS072 e
2		853	99.1	882	21	AAA92031	Staphylococcus aur
3		846.6	98.3	882	23	AAS54848	Staphylococcus aur
4		831.6	96.6	864	23	AAS51602	Staphylococcus aur
c	5	807	93.7	1893	18	AAV74466	Staphylococcus aur
6		449.8	52.2	909	24	ABN90871	Staphylococcus epi
7		449.8	52.2	3444	22	AAH54780	S. epidermidis gen
c	8	449.8	52.2	4045	22	AAH54565	S. epidermidis gen
c	9	413	48.0	413	23	AAS50245	Staphylococcus aur
	10	406.8	47.2	783	22	AAH52415	S. epidermidis ope
c	11	337	39.1	337	23	AAS50698	Staphylococcus aur
	12	252.6	29.3	2944528	24	ABA03041	Listeria monocytog
	13	239.8	27.9	5635	24	ABQ71010	Listeria monocytog
c	14	231.2	26.9	495269	24	ABQ67195	Listeria innocua c
	15	231.2	26.9	3011208	24	ABQ69245	Listeria innocua D
c	16	204	23.7	3454	22	AAH54634	S. epidermidis gen
	17	196.8	22.9	882	23	AAS52896	Enterococcus faeca
	18	195	22.6	7528	20	AAX12992	Enterococcus faeca
	19	175.2	20.3	870	24	ABN66185	Streptococcus poly
	20	174.8	20.3	2365589	24	ABA90521	Genomic sequence o
	21	164	19.0	811	24	ABK73724	Bacillus lichenifo
	22	157.2	18.3	894	14	AAQ39243	FPS DNA. Bacillus
	23	157.2	18.3	894	17	AAT40229	Native farnesyldip
	24	155.8	18.1	1337	20	AAX77873	L. monocytogenes c
	25	155.6	18.1	894	17	AAT40226	Mutant farnesyldip
	26	155.6	18.1	894	17	AAT40227	Mutant farnesyldip
	27	155.6	18.1	894	17	AAT40228	Mutant farnesyldip
	28	155.6	18.1	894	19	AAV38455	DNA encoding farne
	29	155.6	18.1	894	19	AAV18516	Bacillus stearothe
	30	155.6	18.1	894	20	AAX86781	Farnesyl diphospha
	31	155.6	18.1	894	21	AAC63809	Bacillus stearothe
	32	155.6	18.1	894	24	ABK96797	B. stearothermophi
	33	154	17.9	894	17	AAT40225	Mutant farnesyldip
	34	154	17.9	894	20	AAX86780	Geranyl diphosphat
	35	154	17.9	894	24	ABK96808	B. stearothermophi
	36	153.6	17.8	873	24	ABN66184	Streptococcus poly
	37	153.6	17.8	2155561	24	ABN71527	Streptococcus poly
	38	150.6	17.5	486	19	AAV15824	Probe B500 for Mic
	39	137	15.9	1664976	19	AAV21209	Methanococcus jann
	40	136.4	15.8	885	21	AAA13985	Taxus cuspidata ge
	41	134.8	15.7	1179	21	AAA13994	Taxus GGPP synthas
	42	134.8	15.7	1179	21	AAA13995	Taxus GGPP synthas
	43	134.8	15.7	1179	21	AAA13996	Taxus GGPP synthas
	44	134.8	15.7	1179	21	AAA13997	Taxus GGPP synthas
	45	134.8	15.7	1179	21	AAA13998	Taxus GGPP synthas

ALIGNMENTS

RESULT 1
AAS00821
ID AAS00821 standard; DNA; 861 BP.
XX
AC AAS00821;
XX
DT 04-JUL-2001 (first entry)
XX
DE S. aureus HGS072 encoding Farnesyl diphosphatesynthase, IspA.
XX
KW Farnesyl diphosphatesynthase; IspA; immunogen; vaccine; antibody;
KW wound infection; cellulitis; burn infection; eyelid infection;
KW food poisoning; joint infection; neonatal conjunctivitis; osteomyelitis;
KW skin infection; scalded skin syndrome; toxic epidermal necrosis;
KW Ritter's disease; Lyell's disease; toxic shock syndrome; endocarditis;
KW HGS072; ds.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT CDS 1..861
FT /*tag= a
FT /product= "IspA"
FT /partial
FT /note= "No stop codon"
XX
PN WO200116292-A2.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-US23773.
XX

PR 01-SEP-1999; 99US-0151933.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Choi GH;
XX
DR WPI; 2001-183259/18.
DR P-PSDB; AAU00859.
XX
PT New isolated nucleic acid for use in diagnosing Staphylococcus
PT infections and in vaccines for eliciting immune responses to the
PT infections -
XX
PS Claim 1; Page 23; 225pp; English.
XX
CC The sequence encodes S. aureus IspA (Farnesyl diphosphatesynthase).
CC The polynucleotides of the invention are used to detect Staphylococcus
CC nucleic acids in a biological sample from an animal for diagnosing
CC Staphylococcus infections. The polypeptides of the invention are used to
CC detect anti-Staphylococcus antibodies in a biological sample from an
CC animal to diagnose Staphylococcus infections. The polypeptides are also
CC used in vaccines to elicit protective antibodies in an animal to a member
CC of the Staphylococcus genus and for preventing or attenuating an
CC infection caused by a member of the Staphylococcus genus e.g wound
CC infection, cellulitis, burn infection, eyelid infection, food poisoning,
CC joint infection, neonatal conjunctivitis, osteomyelitis, skin infection,
CC scalded skin syndrome (also known as toxic epidermal necrosis, Ritter's
CC disease and Lyell's disease), toxic shock syndrome and endocarditis. The
CC polynucleotides may also be used in vaccines and for preventing or
CC attenuating a Staphylococcus infection. Antibodies to the polypeptides
CC may be used to purify, detect and target the polypeptides in in vitro and
CC in vivo diagnostic and therapeutic methods.
XX
SQ Sequence 861 BP; 323 A; 114 C; 180 G; 244 T; 0 other;

Query Match 100.0%; Score 861; DB 22; Length 861;
Best Local Similarity 100.0%; Pred. No. 3.6e-194;
Matches 861; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||||||||||||||||||||
Db        841 TTATTAGAAATCGTTGATTTA 861
```

RESULT 2
AAA92031
ID AAA92031 standard; DNA; 882 BP.
XX
AC AAA92031;
XX
DT 12-JAN-2001 (first entry)
XX
DE Staphylococcus aureus ispA coding sequence.
XX
KW IspA; bacterial disease; respiratory tract infection;
KW gastrointestinal infection; cardiac infection; Helicobacter pylori;
KW stomach cancer; stomach ulcer; gastritis; ds.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT CDS 1..882
FT /*tag= a
FT /product= "IspA"
XX
PN US6107058-A.
XX
PD 22-AUG-2000.
XX
PF 26-MAR-1999; 99US-0276873.
XX
PR 26-MAR-1999; 99US-0276873.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Gwynn M, Wilding EI;
XX
DR WPI; 2000-578535/54.
DR P-PSDB; AAB23333.
XX
PT Novel farnesyl diphosphate synthase polynucleotide from staphylococcus
PT aureus useful for diagnosis and treatment of bacterial infections and
PT as hybridization probe for isolating genomic clones -
XX
PS Claim 1; column 1-4; 15pp; English.
XX
CC The present sequence is the coding sequence for the Staphylococcus aureus
CC IspA protein. This gene and the protein it encodes can be used in many
CC research assays, as well as treatments for bacterial diseases such as
CC infections of the respiratory tract (including otitis media, bacterial

```
CC   tracheitis, acute epiglottitis, thyroiditis, empyema and lung abscesses),
CC   cardiac infections such as infective endocarditis, gastrointestinal
CC   infections including secretory diarrhoea, splenic abscesses and
CC   retroperitoneal abscesses, CNS infections such as cerebral abscesses, eye
CC   infections (including blepharitis, conjunctivitis, keratitis,
CC   endophthalmitis, preseptal and orbital cellulitis and darcryocystitis),
CC   kidney and urinary tract infections such as epididymitis, intrarenal and
CC   perinephric abscesses and toxic shock syndrome, skin diseases (including
CC   impetigo, folliculitis, cutaneous abscesses, wound infection and
CC   bacterial myositis), bone and joint infections such as septic
CC   arthritis and osteomyelitis, septic thrombophlebitis, food poisoning and
CC   scalded skin syndrome. In addition, they can be used to treat
CC   diseases caused by Helicobacter pylori, including stomach cancer, stomach
CC   ulcers and gastritis.
XX
SQ   Sequence 882 BP; 328 A; 118 C; 186 G; 250 T; 0 other;

  Query Match           99.1%;  Score 853;  DB 21;  Length 882;
  Best Local Similarity   99.4%;  Pred. No. 2.9e-192;
  Matches  856;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGCATGTTGTATTCATTAAATGCT 120

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTGTGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              |||||||||||||||||||||||||||||||||||| | ||||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAAGGAAATTAACAAATCAT 300

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAA 600

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||||||||||||||||||||
Db        841 TTATTAGAAATCGTTGATTTA 861


RESULT 3
AAS54848
ID   AAS54848 standard; DNA; 882 BP.
XX
AC   AAS54848;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus DNA for cellular proliferation protein #1160.
XX
KW   Antisense; ds; prokaryotic cellular proliferation gene;
KW   antibiotic; antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   P-PSDB; AAU36989.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Claim 27; Seq ID No 8485; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence encodes an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 882 BP; 330 A; 116 C; 185 G; 251 T; 0 other;
```

```
  Query Match             98.3%;  Score 846.6;  DB 23;  Length 882;
  Best Local Similarity   99.0%;  Pred. No. 9.3e-191;
  Matches  852;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

          241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCAT 300

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTATTAACATTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAA 600

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTTCAGATTAAAGATGATTTA 660

          661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTAGACTGCTATGGTGATGAAGCGAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATAAAAGTACATACGTGAGTTTATTAGGAAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATAGAGACGCAGCAGTGGATGAACTAATGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||||||||||||||||||||
Db        841 TTATTAGAAATCGTTGATTTA 861


RESULT 4
AAS51602
ID   AAS51602 standard; DNA; 864 BP.
XX
AC   AAS51602;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus DNA for cellular proliferation protein #19.
XX
KW   Antisense; ds; prokaryotic cellular proliferation gene;
KW   antibiotic; antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   P-PSDB; AAU33743.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Claim 27; Seq ID No 4184; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence encodes an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 864 BP; 323 A; 112 C; 182 G; 247 T; 0 other;

  Query Match             96.6%;  Score 831.6;  DB 23;  Length 864;
  Best Local Similarity   98.9%;  Pred. No. 3.3e-187;
  Matches  837;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy         16 ATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTA 60

Qy         76 ATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATC 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATC 120

Qy        136 CGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATG 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATG 180

Qy        196 AAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCA 255
```

```
Db        181 AAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCA 240

Qy        256 GCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAG 315
              ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        241 GCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCATAAAGTATATGGTGAG 300

Qy        316 TGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGT 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGT 360

Qy        376 GATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGT 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGT 420

Qy        436 GGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATT 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATT 480

Qy        496 GATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTTTTGCG 555
              |||||||||||||||||||||||||||||||||||||||| ||||||| ||||||
Db        481 GATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCACTATTAACATTTGCG 540

Qy        556 GTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAGAAAGT 615
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        541 GTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAAGAACATTTAGAAAGT 600

Qy        616 TATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCTATGGT 675
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        601 TATAGTTATCATTTAGGTATGATGTTTCAGATTAAAGATGATTTATTAGACTGCTATGGT 660

Qy        676 GATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTACGTAC 735
              |||||||| |||||||||||||||||||||||||||||||||||||||||||| |||
Db        661 GATGAAGCGAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTACATAC 720

Qy        736 GTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGACGCAGCA 795
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTGAGTTTATTAGGAAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGACGCAGCA 780

Qy        796 GTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACACTTATTAGAAATCGTT 855
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        781 GTGGATGAACTAATGCAAATTGATGAACAATTCAATACAAAACACTTATTAGAAATCGTT 840

Qy        856 GATTTA 861
              ||||||
Db        841 GATTTA 846


RESULT 5
AAV74466/c
ID   AAV74466 standard; DNA; 1893 BP.
XX
AC   AAV74466;
XX
DT   16-MAR-1999  (first entry)
XX
DE   Staphylococcus aureus contig SEQ ID #155.
XX
KW   Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW   cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW   skin infection; surgical wound infection; scalded skin syndrome;
KW   toxic shock syndrome; ds.
XX
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1261..1320
FT                   /*tag= a
FT                   /note= "these bases represent a line of missing text in
FT                          the sequence listing in the specification. They
FT                          are included to maintain the nucleotide numbering
FT                          given in the specification for this DNA sequence"
XX
PN   EP786519-A2.
XX
PD   30-JUL-1997.
XX
PF   07-JAN-1997;   97EP-0100117.
XX
PR   05-JAN-1996;   96US-0009861.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Barash SC,  Choi GH,  Dillon PJ,  Fannon MR,  Kunsch CA;
PI   Rosen CA;
XX
DR   WPI; 1997-374922/35.
XX
PT   Polynucleotide(s) and proteins derived from Staphylococcus aureus -
PT   stored on computer readable medium and used in the production of
PT   anti-S.aureus vaccines
XX
PS   Claim 1; Page 784-785; 3271pp; English.
XX
CC   This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC   of the invention. The DNA sequences are recorded on a computer readable
CC   medium, preferably selected from a floppy or hard disk, random access
CC   memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC   the S.aureus DNA sequences allows putative functions to be assigned so
CC   that protein-encoding or regulatory regions of commercial, therapeutic or
CC   industrial importance can be obtained. Specifically, sequences which are
CC   likely to encode antigens have been identified and these polypeptides can
CC   be used in a vaccine composition against S.aureus infection. The
CC   polypeptides can also be used in a kit for the immunodetection of
CC   S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC   including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC   skin and surgical wound infections, scalded skin syndrome, toxic shock
CC   syndrome, etc. Organisms transformed with the DNA sequences can be used
CC   for recombinant production of the polypeptides. The new DNA sequences
CC   (and their fragments) are useful as primers or probes for isolating
CC   homologues of any of the S.aureus DNA sequences contained on the
CC   computer readable medium.
XX
SQ   Sequence 1893 BP; 552 A; 336 C; 233 G; 708 T; 64 other;

  Query Match             93.7%;  Score 807;  DB 18;  Length 1893;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-181;
  Matches  807;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         55 GTTGCGATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTA 114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GTTGCGATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTA 1201

Qy        115 AATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAAT 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 AATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAAT 1141

Qy        175 ACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCA 234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 ACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCA 1081

Qy        235 CTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACA 294
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACA 1021

Qy        295 AATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAA 354
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 AATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAA 961

Qy        355 GCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTA 414
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 GCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTA 901

Qy        415 CAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATG 474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 CAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATG 841
```

```
Qy        475 CAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACA 534
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 CAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACA 781

Qy        535 GGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACA 594
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACA 721

Qy        595 ACTAAAGAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGAT 654
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACTAAAGAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGAT 661

Qy        655 GATTTATTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTT 714
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 GATTTATTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTT 601

Qy        715 GAAAATAATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTG 774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GAAAATAATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTG 541

Qy        775 ACTTATCATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACA 834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ACTTATCATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACA 481

Qy        835 AAACACTTATTAGAAATCGTTGATTTA 861
              |||||||||||||||||||||||||||
Db        480 AAACACTTATTAGAAATCGTTGATTTA 454
```

RESULT 6
ABN90871
ID ABN90871 standard; DNA; 909 BP.
XX
AC ABN90871;
XX
DT 24-JUL-2002 (first entry)
XX
DE Staphylococcus epidermidis ORF nucleic acid sequence SEQ ID NO:334.
XX
KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW antibacterial; gene therapy; gene; ds.
XX
OS Staphylococcus epidermidis.
XX
PN US6380370-B1.
XX
PD 30-APR-2002.
XX
PF 13-AUG-1998; 98US-0134001.
XX
PR 14-AUG-1997; 97US-055779P.
PR 08-NOV-1997; 97US-064964P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Doucette-Stamm LA, Bush D;
XX
DR WPI; 2002-381255/41.
DR P-PSDB; ABP38326.
XX
PT Novel isolated nucleic acid encoding a Staphylococcus epidermidis
PT polypeptide, useful for diagnosing and treating bacterial infections -
XX
PS Disclosure; SEQ ID 334; 267pp; English.
XX
CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC antibacterial activity and can be used in gene therapy. The sequences
CC can also be used in the diagnosis and treatment of bacterial infections,
CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life
CC cycle or inhibit S. epidermidis infection.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site.
XX
SQ Sequence 909 BP; 361 A; 106 C; 165 G; 277 T; 0 other;

```
  Query Match            52.2%;  Score 449.8;  DB 24;  Length 909;
  Best Local Similarity  70.2%;  Pred. No. 6.2e-97;
  Matches  604;  Conservative    0;  Mismatches  257;  Indels    0;  Gaps    0;

Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              |||| ||| |||| ||||||||||||||||| ||  | | ||||    | |       |
Db         28 ATGAAGAAACTACAGATGAATAAATTAATAAATATAATTAATACATCACTGAATAAGTCA 87

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||| |   ||||  | |  | ||| |  |||||||||||||   ||||||||||||||||
Db         88 ATACAATCATCACCATTAAAAACTAATTTAGAAGAAAGTATGAAATATTCATTAAATGCT 147

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              || |||||  | ||| |||||||  | ||||||  | ||  || |    || || |  ||
Db        148 GGTGGTAAAAGAATCAGACCAGTCATATTATTATTAACACTAAAAATGCTTAACAAAGAT 207

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              ||| |   |||  | || || ||  | ||| | ||||||||||||||| |||||  | |||
Db        208 TATCAACAAGGACTAAATAGTGCTTTAGCATTGGAAATGATTCATACTTATTCTTTAATT 267

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              ||||||||  |||||| |||||||||||| ||||| ||  ||||||||||||||||||||
Db        268 CATGATGATTTACCAGCAATGGATAATGACGATTACCGTAGAGGAAAATTAACAAATCAT 327

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||| |||||||| ||||  || ||  | || |||||||| |||||||| ||||| |||
Db        328 AAAGTTTATGGTGAATGGAAAGCCATTCTTGCTGGTGATGCATTATTAACAAAAGCTTTT 387

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||| |  |||| | |||| ||  |   |||  || ||| |||  || || || |
Db        388 GAATTAGTTTCTAATGATACTACCATTGAAGATAGTGTGAAAGTAAGTATTATAAAAAGA 447

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              || |||| |||||||||| ||| | |||||||| || || ||| |  | |||||| ||||
Db        448 CTTTCAAAAGCAAGTGGACATTTGGGAATGGTGGGTGGCCAAGCGCTTGATATGGAAAGT 507

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              |||||  |  |||||  | | |||||||| |||   || ||  |||| || ||||| ||
Db        508 GAAGGGAAGTCAATTCGTTTAGAAACTTTAGAATCAATTCATGAAACTAAGACAGGCGCT 567

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              ||| ||| |||| | ||||||  ||| | ||| || ||| | || ||  | |  | |
Db        568 TTACTAAATTTTTCAGTTATGGCTGCGGTAGACATTGCTCAAGTAGAACAAAATATTGCT 627

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              |  |||||||    | |||| |||||||||| |||||||| || ||||||||||||||||
Db        628 AAGAATTTAGATGAATTTAGTCATCATTTAGGAATGATGTTTCAAATTAAAGATGATTTA 687

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              | ||    |||||||||||| |||| | || |||||||| ||||| ||| | | ||||
Db        688 CTGGATGTGTATGGTGATGAATCAAAACTTGGCAAAAAAGTAGGCAGTGATATAGTAAAT 747

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              |||||||||| || ||   |||| | || ||||| || |||||||| || || |   ||
Db        748 CATAAAAGTACTTATGTTTCTTTACTTGGAAAAGAAGGAGCAGAAGAAAAGTTAAACAAT 807

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              |||  | |    ||  ||  | |     |||  ||||||  ||| ||||   ||||     |
Db        808 CATCAATATCTTGCTATGAACTGCTTAAATCAAATTTCTGATCAATATGATACTTCTGAA 867

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||   || || || ||||||
Db        868 TTAAGTGATATTGTAGATTTA 888
```

RESULT 7

```
AAH54780
ID   AAH54780 standard; DNA; 3444 BP.
XX
AC   AAH54780;
XX
DT   03-SEP-2001  (first entry)
XX
DE   S. epidermidis genomic polynucleotide sequence SEQ ID NO:4144.
XX
KW   Staphylococcus epidermidis SR1 strain; infection; diagnosis;
KW   vaccination; endocarditis; ds.
XX
OS   Staphylococcus epidermidis.
XX
PN   WO200134809-A2.
XX
PD   17-MAY-2001.
XX
PF   09-NOV-2000; 2000WO-US30782.
XX
PR   09-NOV-1999;   99US-0164258.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Kimmerly WJ;
XX
DR   WPI; 2001-316495/33.
XX
PT   Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT   useful for vaccinating against infections, e.g. endocarditis -
XX
PS   Claim 8; Page 1844-1845; 2188pp; English.
XX
CC   AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC   (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC   (I) and (II) can have antibacterial activity and therefore can be used
CC   in vaccination. The nucleic acids (I) may be used to produce the
CC   S. epidermidis polypeptides (II) via the production of vectors
CC   containing them which are used to produce hosts cells which express the
CC   polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC   used to vaccinate subjects and to raise antibodies against the bacteria.
CC   The polypeptides may also be used to assay for other inhibitors of their
CC   activity and therefore identify compounds that may be used for the
CC   treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC   AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC   polynucleotide sequences from the present invention. AAH55091 to
CC   AAH55098 represent oligonucleotide sequences and primers which are used
CC   in the exemplification of the present invention.
CC   N.B. The present invention specifically claims all the polynucleotide
CC   sequences given in the sequence listing of the present specification,
CC   however the sequence listing only goes up to SEQ ID NO:4454 so even
CC   though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC   no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ   Sequence 3444 BP; 1365 A; 451 C; 571 G; 1057 T; 0 other;

  Query Match           52.2%;  Score 449.8;  DB 22;  Length 3444;
  Best Local Similarity  70.2%;  Pred. No. 8.4e-97;
  Matches  604;  Conservative    0;  Mismatches  257;  Indels    0;  Gaps    0;

Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              |||| ||| |||| ||||||||||||||||| ||  | | ||||    | |       |
Db        766 ATGAAGAAACTACAGATGAATAAATTAATAAATATAATTAATACATCACTGAATAAGTCA 825

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||| |   ||||  | |  | ||| |  ||||||||||||||   ||||||||||||||
Db        826 ATACAATCATCACCATTAAAAACTAATTTAGAAGAAAGTATGAAATATTCATTAAATGCT 885

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              || |||||| | ||| |||||||  | |||||| | ||  || |    || || |  ||
Db        886 GGTGGTAAAAGAATCAGACCAGTCATATTATTATTAACACTAAAAATGCTTAACAAAGAT 945

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              ||| |   |||  | || || ||  | ||| | ||||||||||||||| |||||  | |||
Db        946 TATCAACAAGGACTAAATAGTGCTTTAGCATTGGAAATGATTCATACTTATTCTTTAATT 1005

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              ||||||||  |||||||  |||||||||||| ||||| ||  ||||||||||||||||||
Db       1006 CATGATGATTTACCAGCAATGGATAATGACGATTACCGTAGAGGAAAATTAACAAATCAT 1065

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||| ||||||||  ||||  || ||  | || ||||||||  ||||||||  ||||| |||
Db       1066 AAAGTTTATGGTGAATGGAAAGCCATTCTTGCTGGTGATGCATTATTAACAAAAGCTTTT 1125

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||| |  |||| | ||||  ||   |    |||   || ||| |||   || || || |
Db       1126 GAATTAGTTTCTAATGATACTACCATTGAAGATAGTGTGAAAGTAAGTATTATAAAAAGA 1185

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              || |||| |||||||||  ||| | ||||||||  || ||  |||  | | ||||| ||||
Db       1186 CTTTCAAAAGCAAGTGGACATTTGGGAATGGTGGGTGGCCAAGCGCTTGATATGGAAAGT 1245

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              |||||  |  |||||  | | ||||||||  |||   || ||  |||| || ||||| ||
Db       1246 GAAGGGAAGTCAATTCGTTTAGAAACTTTAGAATCAATTCATGAAACTAAGACAGGCGCT 1305

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              ||| ||| |||| | ||||||  ||| | ||| || ||| | || ||  | |  | |
Db       1306 TTACTAAATTTTTCAGTTATGGCTGCGGTAGACATTGCTCAAGTAGAACAAAATATTGCT 1365

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              |  |||||||    | |||| ||||||||||  |||||||| || ||||||||||||||
Db       1366 AAGAATTTAGATGAATTTAGTCATCATTTAGGAATGATGTTTCAAATTAAAGATGATTTA 1425

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              | ||    |||||||||||| ||||  | ||  |||||||| |||||  ||| | | ||||
Db       1426 CTGGATGTGTATGGTGATGAATCAAAACTTGGCAAAAAAGTAGGCAGTGATATAGTAAAT 1485

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||||  || ||   |||| | || ||||| || ||||||||  || || |   ||
Db       1486 CATAAAAGTACTTATGTTTCTTTACTTGGAAAAGAAGGAGCAGAAGAAAAGTTAAACAAT 1545

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              |||  | |   ||  || |    |||  ||||||  |||  ||||   ||||     |
Db       1546 CATCAATATCTTGCTATGAACTGCTTAAATCAAATTTCTGATCAATATGATACTTCTGAA 1605

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||   || || || ||||||
Db       1606 TTAAGTGATATTGTAGATTTA 1626


RESULT 8
AAH54565/c
ID   AAH54565 standard; DNA; 4045 BP.
XX
AC   AAH54565;
XX
DT   03-SEP-2001  (first entry)
XX
DE   S. epidermidis genomic polynucleotide sequence SEQ ID NO:3929.
XX
KW   Staphylococcus epidermidis SR1 strain; infection; diagnosis;
KW   vaccination; endocarditis; ds.
XX
OS   Staphylococcus epidermidis.
XX
PN   WO200134809-A2.
XX
PD   17-MAY-2001.
XX
PF   09-NOV-2000; 2000WO-US30782.
XX
PR   09-NOV-1999;   99US-0164258.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
```

```
PI   Kimmerly WJ;
XX
DR   WPI; 2001-316495/33.
XX
PT   Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT   useful for vaccinating against infections, e.g. endocarditis -
XX
PS   Claim 8; Page 1586-1588; 2188pp; English.
XX
CC   AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC   (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC   (I) and (II) can have antibacterial activity and therefore can be used
CC   in vaccination. The nucleic acids (I) may be used to produce the
CC   S. epidermidis polypeptides (II) via the production of vectors
CC   containing them which are used to produce hosts cells which express the
CC   polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC   used to vaccinate subjects and to raise antibodies against the bacteria.
CC   The polypeptides may also be used to assay for other inhibitors of their
CC   activity and therefore identify compounds that may be used for the
CC   treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC   AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC   polynucleotide sequences from the present invention. AAH55091 to
CC   AAH55098 represent oligonucleotide sequences and primers which are used
CC   in the exemplification of the present invention.
CC   N.B. The present invention specifically claims all the polynucleotide
CC   sequences given in the sequence listing of the present specification,
CC   however the sequence listing only goes up to SEQ ID NO:4454 so even
CC   though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC   no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ   Sequence 4045 BP; 1249 A; 691 C; 536 G; 1569 T; 0 other;

  Query Match           52.2%;  Score 449.8;  DB 22;  Length 4045;
  Best Local Similarity  70.2%;  Pred. No. 8.7e-97;
  Matches  604;  Conservative    0;  Mismatches  257;  Indels    0;  Gaps    0;

Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              |||| ||| |||| ||||||||||||||||| ||  | | ||||    | |         |
Db       2313 ATGAAGAAACTACAGATGAATAAATTAATAAATATAATTAATACATCACTGAATAAGTCA 2254

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||| |   ||||  | |  | ||| |  ||||||||||||||   ||||||||||||||
Db       2253 ATACAATCATCACCATTAAAAACTAATTTAGAAGAAAGTATGAAATATTCATTAAATGCT 2194

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              || |||||| | ||| |||||||  | |||||| | ||  || |    || || |  ||
Db       2193 GGTGGTAAAAGAATCAGACCAGTCATATTATTATTAACACTAAAAATGCTTAACAAAGAT 2134

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              ||| |   |||  | || || ||  | ||| | ||||||||||||||| |||||  | |||
Db       2133 TATCAACAAGGACTAAATAGTGCTTTAGCATTGGAAATGATTCATACTTATTCTTTAATT 2074

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              ||||||||  |||||||  |||||||||||  |||||  ||  ||||||||||||||||||
Db       2073 CATGATGATTTACCAGCAATGGATAATGACGATTACCGTAGAGGAAAATTAACAAATCAT 2014

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              |||||  ||||||||  ||||   ||  ||   | ||  ||||||||  ||||||||  ||||| |||
Db       2013 AAAGTTTATGGTGAATGGAAAGCCATTCTTGCTGGTGATGCATTATTAACAAAAGCTTTT 1954

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              |||  |  |||| | ||||   ||    |    |||    || ||| |||    || || || |
Db       1953 GAATTAGTTTCTAATGATACTACCATTGAAGATAGTGTGAAAGTAAGTATTATAAAAAGA 1894

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              || |||| |||||||||  ||| | ||||||||  || ||  |||   | |  ||||||  ||||
Db       1893 CTTTCAAAAGCAAGTGGACATTTGGGAATGGTGGGTGGCCAAGCGCTTGATATGGAAAGT 1834

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              |||||  |  |||||  | | ||||||||  |||   ||  ||  ||||  || |||||  ||
Db       1833 GAAGGGAAGTCAATTCGTTTAGAAACTTTAGAATCAATTCATGAAACTAAGACAGGCGCT 1774

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
```

```
              ||| ||| |||| | ||||||  ||| | ||| || ||| | || ||  | |  | |
Db       1773 TTACTAAATTTTTCAGTTATGGCTGCGGTAGACATTGCTCAAGTAGAACAAAATATTGCT 1714

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
               |  |||||||  | |||| ||||||||| |||||||| || ||||||||||||||||
Db       1713 AAGAATTTAGATGAATTTAGTCATCATTTAGGAATGATGTTTCAAATTAAAGATGATTTA 1654

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
               | ||    |||||||||||| ||||  | || |||||||| ||||| ||| | | ||||
Db       1653 CTGGATGTGTATGGTGATGAATCAAAACTTGGCAAAAAAGTAGGCAGTGATATAGTAAAT 1594

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||||  || ||   |||| | || ||||| || |||||||| || || |   ||
Db       1593 CATAAAAGTACTTATGTTTCTTTACTTGGAAAAGAAGGAGCAGAAGAAAAGTTAAACAAT 1534

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              |||  | |    ||  || |     |||  ||||||  ||| ||||   ||||     |
Db       1533 CATCAATATCTTGCTATGAACTGCTTAAATCAAATTTCTGATCAATATGATACTTCTGAA 1474

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||   || || || ||||||
Db       1473 TTAAGTGATATTGTAGATTTA 1453


RESULT 9
AAS50245/c
ID   AAS50245 standard; DNA; 413 BP.
XX
AC   AAS50245;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus cellular proliferation inhibitory sequence #1469.
XX
KW   Antisense; ss; prokaryotic cellular proliferation;
KW   antibiotic; antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Claim 1; Seq ID No 2822; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
```

```
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence is an antisense
CC   oligonucleotide of the invention.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 413 BP; 115 A; 89 C; 58 G; 151 T; 0 other;

  Query Match             48.0%;  Score 413;  DB 23;  Length 413;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-88;
  Matches  413;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        160 TTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATG 219
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 TTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATG 354

Qy        220 ATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGA 279
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 ATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGA 294

Qy        280 CGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGAT 339
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 CGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGAT 234

Qy        340 GCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTA 399
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 GCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTA 174

Qy        400 AAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGT 459
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 AAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGT 114

Qy        460 CAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATA 519
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 CAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATA 54

Qy        520 CACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCAGA 572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         53 CACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCAGA 1


RESULT 10
AAH52415
ID   AAH52415 standard; DNA; 783 BP.
XX
AC   AAH52415;
XX
DT   03-SEP-2001  (first entry)
XX
DE   S. epidermidis open reading frame nucleotide sequence SEQ ID NO:223.
XX
KW   Staphylococcus epidermidis SR1 strain; infection; diagnosis;
KW   vaccination; endocarditis; ds.
XX
OS   Staphylococcus epidermidis.
XX
PN   WO200134809-A2..
XX
PD   17-MAY-2001.
XX
PF   09-NOV-2000; 2000WO-US30782.
XX
PR   09-NOV-1999;   99US-0164258.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Kimmerly WJ;
XX
DR   WPI; 2001-316495/33.
DR   P-PSDB; AAG81565.
XX
PT   Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT   useful for vaccinating against infections, e.g. endocarditis -
XX
PS   Claim 8; Page 103; 2188pp; English.
XX
CC   AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC   (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC   (I) and (II) can have antibacterial activity and therefore can be used
CC   in vaccination. The nucleic acids (I) may be used to produce the
CC   S. epidermidis polypeptides (II) via the production of vectors
CC   containing them which are used to produce hosts cells which express the
CC   polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC   used to vaccinate subjects and to raise antibodies against the bacteria.
CC   The polypeptides may also be used to assay for other inhibitors of their
CC   activity and therefore identify compounds that may be used for the
CC   treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC   AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC   polynucleotide sequences from the present invention. AAH55091 to
CC   AAH55098 represent oligonucleotide sequences and primers which are used
CC   in the exemplification of the present invention.
CC   N.B. The present invention specifically claims all the polynucleotide
CC   sequences given in the sequence listing of the present specification,
CC   however the sequence listing only goes up to SEQ ID NO:4454 so even
CC   though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC   no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ   Sequence 783 BP; 297 A; 93 C; 147 G; 246 T; 0 other;

  Query Match             47.2%;  Score 406.8;  DB 22;  Length 783;
  Best Local Similarity   70.9%;  Pred. No. 8.8e-87;
  Matches  540;  Conservative    0;  Mismatches  222;  Indels    0;  Gaps    0;

Qy        100 ATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACT 159
              |||   |||||||||||||||| |||||| | ||| |||||||  | |||||| | ||
Db          1 ATGAAATATTCATTAAATGCTGGTGGTAAAAGAATCAGACCAGTCATATTATTATTAACA 60

Qy        160 TTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATG 219
              || |    || || |  || ||| |   |||  | || || ||  | ||| | ||||||
Db         61 CTAAAAATGCTTAACAAAGATTATCAACAAGGACTAAATAGTGCTTTAGCATTGGAAATG 120

Qy        220 ATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGA 279
              |||||||| |||||  | ||||||||||  |||||||  |||||||||| ||||| ||
Db        121 ATTCATACTTATTCTTTAATTCATGATGATTTACCAGCAATGGATAATGACGATTACCGT 180

Qy        280 CGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGAT 339
              |||||||||||||||||||||||| |||||||| ||||  || ||  | || ||||||
Db        181 AGAGGAAAATTAACAAATCATAAAGTTTATGGTGAATGGAAAGCCATTCTTGCTGGTGAT 240

Qy        340 GCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTA 399
              || |||||||| ||||| |||||| |  |||| | ||||  ||   |    |||   ||
Db        241 GCATTATTAACAAAAGCTTTTGAATTAGTTTCTAATGATACTACCATTGAAGATAGTGTG 300

Qy        400 AAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGT 459
              ||| |||   || || || | || |||| |||||||||| ||| | |||||||| || ||
Db        301 AAAGTAAGTATTATAAAAAGACTTTCAAAAGCAAGTGGACATTTGGGAATGGTGGGTGGC 360

Qy        460 CAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATA 519
              |||  | | |||||| |||| |||||  |  |||||  | | |||||||| |||   ||
Db        361 CAAGCGCTTGATATGGAAAGTGAAGGGAAGTCAATTCGTTTAGAAACTTTAGAATCAATT 420

Qy        520 CACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCT 579
              ||  |||| || ||||| || ||| ||| |||| | ||||||  ||| | ||| || |||
Db        421 CATGAAACTAAGACAGGCGCTTTACTAAATTTTTCAGTTATGGCTGCGGTAGACATTGCT 480

Qy        580 AATGTCGATGATACAACTAAAGAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATG 639
              | || ||  | |  | |    |  |||||||    | |||| ||||||||| |||||
Db        481 CAAGTAGAACAAAATATTGCTAAGAATTTAGATGAATTTAGTCATCATTTAGGAATGATG 540

Qy        640 TTCCAGATTAAAGATGATTTATTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAA 699
```

```
                || || ||||||||||||||| | ||    |||||||||||| ||||  | || ||||||
Db        541 TTTCAAATTAAAGATGATTTACTGGATGTGTATGGTGATGAATCAAAACTTGGCAAAAAA 600

Qy        700 GTGGGCAGCGATCTTGAAAATAATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGC 759
                || ||||| ||| | | |||| |||||||||| || ||   |||| | || ||||| ||
Db        601 GTAGGCAGTGATATAGTAAATCATAAAAGTACTTATGTTTCTTTACTTGGAAAAGAAGGA 660

Qy        760 GCAGAAGATAAATTGACTTATCATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGAT 819
                |||||||| || || |   ||||| | |    ||  || |     |||  |||||| |
Db        661 GCAGAAGAAAAGTTAAACAATCATCAATATCTTGCTATGAACTGCTTAAATCAAATTTCT 720

Qy        820 GAACAATTCAATACAAAACACTTATTAGAAATCGTTGATTTA 861
                || ||||   ||||     | |||   || || || ||||||
Db        721 GATCAATATGATACTTCTGAATTAAGTGATATTGTAGATTTA 762
```

RESULT 11
AAS50698/c
ID AAS50698 standard; DNA; 337 BP.
XX
AC AAS50698;
XX
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation inhibitory sequence #1922.
XX
KW Antisense; ss; prokaryotic cellular proliferation;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Claim 1; Seq ID No 3275; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence is an antisense
CC oligonucleotide of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 337 BP; 96 A; 71 C; 47 G; 123 T; 0 other;

```
  Query Match             39.1%;  Score 337;  DB 23;  Length 337;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-70;
  Matches  337;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        225 TACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGG 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 TACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGG 278

Qy        285 AAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTT 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 AAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTT 218

Qy        345 ATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAAT 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 ATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAAT 158

Qy        405 AAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAAT 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 AAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAAT 98

Qy        465 GTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAA 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 GTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAA 38

Qy        525 AACAAAAACAGGAGCATTATTAACTTTTGCGGTTATG 561
              |||||||||||||||||||||||||||||||||||||
Db         37 AACAAAAACAGGAGCATTATTAACTTTTGCGGTTATG 1
```

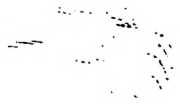
RESULT 12
ABA03041
ID ABA03041 standard; DNA; 2944528 BP.
XX
AC ABA03041;
XX
DT 05-FEB-2002 (first entry)
XX
DE Listeria monocytogenes EGD-e genome sequence.
XX
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease; ds.
XX
OS Listeria monocytogenes.
XX
PN WO200177335-A2.
XX
PD 18-OCT-2001.
XX
PF 11-APR-2001; 2001WO-FR01118.
XX
PR 11-APR-2000; 2000FR-0004629.
XX
PA (INSP) INST PASTEUR.
XX
PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
PI Rose M, Voss H;
XX
DR WPI; 2002-010914/01.
XX
PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT and prevention of Listeria and related bacterial infections, and

```
PT   related polypeptides -
XX
PS   Claim 1; SEQ ID No 1; 192pp; French.
XX
CC   The present sequence is the genome sequence of Listeria monocytogenes
CC   EGD-e. This sequence and fragments of this sequence are useful for
CC   selecting probes and primers for detecting genes in L. monocytogenes and
CC   related organisms, and to study genetic polymorphisms and other genomes.
CC   Proteins (ABB47297-ABB50149) expressed from the present sequence are
CC   useful for raising specific antibodies, identification of L.
CC   monocytogenes and related organisms, and for biosynthesis and
CC   biodegradation, especially biosynthesis of Vitamin B12. This sequence and
CC   proteins encoded by it are also useful for selecting compounds that
CC   regulate gene expression and cell replication and modulate L.
CC   monocytogenes-related diseases. In addition, this sequence and proteins
CC   encoded by it are useful in pharmaceutical and vaccines compositions for
CC   the treatment or prevention of infections by L. monocytogenes and related
CC   organisms.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.

SQ   Sequence 2944528 BP; 914202 A; 563301 C; 555061 G; 911964 T; 0 other;

  Query Match             29.3%;  Score 252.6;  DB 24;  Length 2944528;
  Best Local Similarity   59.3%;  Pred. No. 1.6e-49;
  Matches  429;  Conservative    0;  Mismatches  294;  Indels    0;  Gaps    0;

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              |||||  |   |   || ||  ||   ||| ||||   |||||  |||||  |  | ||
Db    1386000 ATAAACGAGCGAAATATCGAACCTAGACTAAAAGAGTCCATGTTATATTCTATTCAAGCA 1386059

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              || || |||||  || || ||| | ||  | ||    ||  |  |  |  ||||     |
Db    1386060 GGCGGGAAACGGATTCGTCCAATGCTAGTTTTCGCTACGCTTCAAGCTTTAAAAGTAAAT 1386119

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
                 |||||| | || | |||||  ||  |||||||||||||| |||||    ||||||
Db    1386120 CCGCTTTTAGGTGTAAAAACCGCAACAGCGTTAGAAATGATTCACACATACAGCCTTATT 1386179

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              |||||||| |||||||| |||||||| |||||||||||| ||||| || |  || ||||||
Db    1386180 CATGATGATCTACCAGCAATGGATAACGATGATTATCGTCGAGGCAAGTGGACTAATCAT 1386239

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||| | ||| ||    |||||||| || ||||| |||||||| ||||    || |||
Db    1386240 AAAGTTTTTGGCGATGCAACTGCGATTTTGGCAGGAGATGCTTTACTAACGCTCGCTTTT 1386299

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
                 || |  |    || ||||  ||| ||  |||   |   ||     |  |  | |
Db    1386300 TCTATTTTAGCTGAAGACGATAATTTATCTTTTGAGACACGCATTGCTTTGATTAACCAA 1386359

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              |    |    || |||   |  |||||||| || ||||||  |   |||  || ||
Db    1386360 ATTAGTTTTAGTAGCGGTGCAGAAGGAATGGTTGGTGGTCAACTTGCAGACTTGGAAGCG 1386419

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              |||  | ||| | |    || |||   ||    |  || ||    |  |||||| || | |
Db    1386420 GAAAACAAACAAGTGACGCTAGAAGAGTTATCATCCATTCATGCACGAAAAACGGGTGAA 1386479

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              ||||||| || ||| || |    |||||||  | || ||  | |  |||     |   |
Db    1386480 TTATTAATTTATGCTGTAACCTCTGCAGCAAAAATTGCGGAAGCTGATCCAGAACAAACG 1386539

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              |||  |||  || || |    |  || | || ||    || || ||||  || ||| |
Db    1386540 AAACGCTTACGAATTTTTGCAGAGAATATTGGGATTGGATTTCAAATTAGCGACGATATT 1386599

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              |||||      |||||||||| | ||  | ||||||||    ||   |||   |
Db    1386600 TTAGATGTAATTGGTGATGAAACGAAAATGGGTAAAAAGACAGGGGCCGACGCTTTTCTG 1386659

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              |||||||||||| ||     | ||| |   |   ||||| ||| ||     ||| | | |
Db    1386660 AATAAAAGTACCTATCCCGGATTACTCACGCTTGATGGGGCAAAAAGGGCATTAAATGAG 1386719

Qy        781 CAT 783
              |||
Db    1386720 CAT 1386722


RESULT 13
ABQ71010
ID   ABQ71010 standard; DNA; 5635 BP.
XX
AC   ABQ71010;
XX
DT   29-AUG-2002  (first entry)
XX
DE   Listeria monocytogenes 4b contig DNA sequence #952.
XX
KW   Antibacterial; Listeria; food contamination; mutational analysis;
KW   infection; ds.
XX
OS   Listeria monocytogenes 4b.
XX
PN   WO200228891-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-FR03061.
XX
PR   04-OCT-2000; 2000FR-0012697.
XX
PA   (INSP ) INST PASTEUR.
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Kunst F,  Glaser P;
XX
DR   WPI; 2002-332479/37.
XX
PT   New genomic sequences from Listeria species, useful for detection,
PT   treatment and prevention of infection, also related polypeptides,
PT   antibodies and modulators -
XX
PS   Claim 14; SEQ ID 3823; 180pp; French.
XX
CC   The present invention relates to nucleic acid sequences
CC   (ABQ67188-ABQ71212) from Listeria sp. The sequences are useful as probes
CC   and primers for identification and/or detection of Listeria (e.g. as
CC   contaminants in foods, or mutational analysis) and for analysis of
CC   gene expression. Proteins encoded by the nucleic acid sequences can be
CC   used to screen for compounds that modulate gene expression, replication
CC   and pathogenicity of Listeria (potential therapeutic agents), also for
CC   treating infections by Listeria, and are useful as immunogens in
CC   anti-Listeria vaccines.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 5635 BP; 1896 A; 927 C; 1238 G; 1574 T; 0 other;

  Query Match             27.9%;  Score 239.8;  DB 24;  Length 5635;
  Best Local Similarity   58.2%;  Pred. No. 4.2e-47;
  Matches  421;  Conservative    0;  Mismatches  302;  Indels    0;  Gaps    0;

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||||   |   |   || ||  ||   ||| ||||    |||||  |||||  |  | ||
Db       2944 ATAAGCGAGCGAAATATCGAACCTAGACTAAAAGAGTCCATGTTATATTCTGTTCAAGCA 3003

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              || || |||||  || || ||  | ||  | ||    ||  |  |  |  ||||     |
Db       3004 GGTGGAAAACGGATTCGTCCCATGCTAGTTTTTGCTACACTTCAAGCCTTAAAAGTAAAT 3063

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
                 |||||| |||| |  ||||  ||  ||||||||||||||| |||   || ||
```

```
Db       3064 CCGCTTTTAGGTGTGAAAACAGCAACGGCGTTAGAAATGATTCATACGTATAGCCTGATC 3123

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              |||||||| |||||| |||||||||||||| ||||| || || |||| || ||||||
Db       3124 CATGATGATTTACCAGCAATGGATAATGATGACTATCGTCGCGGCAAATATACTAATCAT 3183

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||| | ||| ||    |||||||| || ||||| || |||||| |||| || |||
Db       3184 AAAGTTTTTGGGGATGCAACTGCGATTTTGGCAGGAGACGCTTTACTAACGCTTGCTTTT 3243

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
                || |  |    || || |  ||| ||  |||| |   |||    |  |  | |
Db       3244 TCTATTTTAGCGGAAGACGAAAATTTATCTTTTGAAACACGCATAGCTTTGATTAATCAA 3303

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
               |    |    ||||||   |  |||||||| |||||||| |   |||||||  |
Db       3304 ATTAGTTTTAGTAGTGGTGCAGAAGGAATGGTTGGCGGTCAACTTGCAGATATGGAGGCA 3363

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              |||  | ||| | |    | |||   |  |   || ||  |  ||||||||  | |
Db       3364 GAAAACAAACAAGTGACATTAGAAGAACTATCATCCATCCATGCACGGAAAACAGGCGAA 3423

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              ||||||| |||||| || |    ||| ||  | || ||  | |  ||      |  |
Db       3424 TTATTAATTTTTGCCGTAACTTCTGCTGCGAAAATTGCAGAAGCTGACCCGGAACAAACG 3483

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
               ||  |||   | || |    |  || | || ||    ||||| || |  || ||| |
Db       3484 AAAAGATTACGGATTTTTGCGGAAAATATTGGGATTGGATTCCAAATCAGCGACGATATT 3543

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              |||||     |||||||||| ||||  | ||||||||    ||   |||   |
Db       3544 TTAGATGTTATTGGTGATGAAACAAAAATGGGTAAAAAGACAGGGGTCGACGCTTTTCTG 3603

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              || |||||||| ||     | ||| |   |   |||| || ||     ||| | | |
Db       3604 AACAAAAGTACCTATCCCGGATTACTCACGCTTGATGGTGCCAAAAGGGCATTAAATGAG 3663

Qy        781 CAT 783
              |||
Db       3664 CAT 3666


RESULT 14
ABQ67195/c
ID   ABQ67195 standard; DNA; 495269 BP.
XX
AC   ABQ67195;
XX
DT   29-AUG-2002  (first entry)
XX
DE   Listeria innocua contig DNA sequence #8.
XX
KW   Antibacterial; Listeria; food contamination; mutational analysis;
KW   infection; ds.
XX
OS   Listeria innocua.
XX
PN   WO200228891-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-FR03061.
XX
PR   04-OCT-2000; 2000FR-0012697.
XX
PA   (INSP ) INST PASTEUR.
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Kunst F,  Glaser P;
XX
DR   WPI; 2002-332479/37.
XX
PT   New genomic sequences from Listeria species, useful for detection,
PT   treatment and prevention of infection, also related polypeptides,
PT   antibodies and modulators -
XX
PS   Claim 5; SEQ ID 8; 180pp; French.
XX
CC   The present invention relates to nucleic acid sequences
CC   (ABQ67188-ABQ71212) from Listeria sp. The sequences are useful as probes
CC   and primers for identification and/or detection of Listeria (e.g. as
CC   contaminants in foods, or mutational analysis) and for analysis of
CC   gene expression. Proteins encoded by the nucleic acid sequences can be
CC   used to screen for compounds that modulate gene expression, replication
CC   and pathogenicity of Listeria (potential therapeutic agents), also for
CC   treating infections by Listeria, and are useful as immunogens in
CC   anti-Listeria vaccines.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 495269 BP; 159529 A; 88849 C; 93135 G; 153754 T; 2 other;

  Query Match             26.9%;  Score 231.2;  DB 24;  Length 495269;
  Best Local Similarity   56.4%;  Pred. No. 1.2e-44;
  Matches  431;  Conservative    0;  Mismatches  333;  Indels    0;  Gaps    0;

Qy         20 ATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTAATGG 79
              || | |  | | | | | |  || |     | |     |  |||||||         || |
Db     366205 ATCATTATAAAAAAGTACTTGATGAGTCGCTTTTTAAAGAAATAAATATGCGCAATATCG 366146

Qy         80 ATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCGAC 139
              |  || || |  |||||    ||||| |||||| |  | |||||  || ||||| || || |
Db     366145 AACCTAAGTTGAAAGAATCAATGTTATATTCAGTGCAAGCTGGCGGAAAACGAATTCGTC 366086

Qy        140 CAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGA 199
              || | ||  | ||    ||  |  |  |  ||||||  |||     |   ||| | || |
Db     366085 CAATGCTTGTTTTTGCAACGCTTCAAGCCTTAAATATTGAGCCAATGCGCGGTTTAAAAA 366026

Qy        200 GCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAGCGA 259
              || |  ||||| ||||||||||||||| ||     | ||||| |||||  |||||||
Db     366025 CAGCTACGGCACTTGAAATGATTCATACGTACAGCTTAATTCACGATGATTTACCAGCAA 365966

Qy        260 TGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGA 319
              ||||| |||||||  |||||||| || ||||  || || |||||| |||| |||||
Db     365965 TGGATGATGATGACTATCGACGCGGGAAATGGACGAACCATAAAATATACGGTGATGCAA 365906

Qy        320 CTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATG 379
              | || || |||||||| |||||||| |||||    || |||    || |  |    ||||
Db     365905 CAGCAATTTTAGCAGGAGATGCTTTGTTAACACTCGCTTTTTCTATTTTAGCTGAAGATG 365846

Qy        380 ATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTC 439
              | |   || ||   |||      ||     |  | | |   |          |||||
Db     365845 AAAATCTATCTTTCGAAACGCGTATTGCTTTAATTAATCAAATTAGTTATAGCAGTGGAG 365786

Qy        440 ATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATC 499
              |  || ||||| || || |||      ||||||  ||   |||    ||| | |
Db     365785 CAGAAGGCATGGTAGGTGGGCAACAAGCCGATATGGAAGCTGAAAATAAACAAGTCACAT 365726

Qy        500 TTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTA 559
              | |||   | | |   || |||      ||||| || | | |||||| ||| || || |
Db     365725 TAGAAGAACTAGCATCAATCCACGCTCGCAAAACTGGTGAACTATTAATTTTCGCAGTAA 365666

Qy        560 TGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAGAAAGTTATA 619
                 || ||| | |||||| | |    |      |   | | |||  |||  || || |
Db     365665 CCTCAGCCGCAAAAATCGCTGAAGCAACTCCAGAACAAACAAAACGATTACGAATTTTTG 365606

Qy        620 GTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCTATGGTGATG 679
              |  || | || ||    || || ||||  || ||| | |||||      ||| ||||
Db     365605 CAGAAAATATCGGCATTGGTTTTCAAATTAGCGACGATATTTTAGATGTAATTGGCGATG 365546

Qy        680 AAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTACGTACGTGA 739
              || |||| | ||||||||    ||  |||  |      ||||||||||| ||
Db     365545 AAACAAAAATGGGTAAAAAGACAGGGGTCGACGCTTTTCTGAATAAAAGTACCTATCCCG 365486
```

```
Qy        740 GTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTATCAT 783
              | ||| |       || || ||  ||     ||| | | | |||
Db     365485 GATTACTCACACTCGAAGGTGCTAAACGGGCATTAAATGAGCAT 365442


RESULT 15
ABQ69245
ID   ABQ69245 standard; DNA; 3011208 BP.
XX
AC   ABQ69245;
XX
DT   29-AUG-2002  (first entry)
XX
DE   Listeria innocua DNA sequence #684.
XX
KW   Antibacterial; Listeria; food contamination; mutational analysis;
KW   infection; ds.
XX
OS   Listeria innocua.
XX
PN   WO200228891-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-FR03061.
XX
PR   04-OCT-2000; 2000FR-0012697.
XX
PA   (INSP ) INST PASTEUR.
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Kunst F,  Glaser P;
XX
DR   WPI; 2002-332479/37.
XX
PT   New genomic sequences from Listeria species, useful for detection,
PT   treatment and prevention of infection, also related polypeptides,
PT   antibodies and modulators  -
XX
PS   Claim 5; SEQ ID 2058; 180pp; French.
XX
CC   The present invention relates to nucleic acid sequences
CC   (ABQ67188-ABQ71212) from Listeria sp. The sequences are useful as probes
CC   and primers for identification and/or detection of Listeria (e.g. as
CC   contaminants in foods, or mutational analysis) and for analysis of
CC   gene expression. Proteins encoded by the nucleic acid sequences can be
CC   used to screen for compounds that modulate gene expression, replication
CC   and pathogenicity of Listeria (potential therapeutic agents), also for
CC   treating infections by Listeria, and are useful as immunogens in
CC   anti-Listeria vaccines.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 3011208 BP; 941651 A; 568176 C; 559189 G; 942192 T; 0 other;

  Query Match           26.9%;  Score 231.2;  DB 24;  Length 3011208;
  Best Local Similarity   56.4%;  Pred. No. 1.9e-44;
  Matches  431;  Conservative    0;  Mismatches  333;  Indels    0;  Gaps    0;

Qy         20 ATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTAATGG 79
              || | |  | | | | | | || |     | |     |  |||||||        || |
Db    1392746 ATCATTATAAAAAAGTACTTGATGAGTCGCTTTTTAAAGAAATAAATATGCGCAATATCG 1392805

Qy         80 ATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCGAC 139
              |  || || |  |||||   ||||| |||||| |  | ||||| || |||||| || || |
Db    1392806 AACCTAAGTTGAAAGAATCAATGTTATATTCAGTGCAAGCTGGCGGAAAACGAATTCGTC 1392865

Qy        140 CAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGA 199
              || | ||  | ||    ||  | |  | ||||||  |||     |   ||| | || |
Db    1392866 CAATGCTTGTTTTTGCAACGCTTCAAGCCTTAAATATTGAGCCAATGCGCGGTTTAAAAA 1392925

Qy        200 GCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAGCGA 259
```

```
                || |  ||||| |||||||||||||| ||     | ||||| |||||  ||||||| |
Db    1392926 CAGCTACGGCACTTGAAATGATTCATACGTACAGCTTAATTCACGATGATTTACCAGCAA 1392985

Qy        260 TGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGA 319
              ||||| ||||||| |||||||| || ||||  || || |||||| |||| |||||    |
Db    1392986 TGGATGATGATGACTATCGACGCGGGAAATGGACGAACCATAAAATATACGGTGATGCAA 1393045

Qy        320 CTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATG 379
              | || || |||||||| |||||||| |||||     || |||     |  |  |    ||||
Db    1393046 CAGCAATTTTAGCAGGAGATGCTTTGTTAACACTCGCTTTTTCTATTTTAGCTGAAGATG 1393105

Qy        380 ATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTC 439
              | |   || ||   |||       ||      |  |  | |   |          |||||
Db    1393106 AAAATCTATCTTTCGAAACGCGTATTGCTTTAATTAATCAAATTAGTTATAGCAGTGGAG 1393165

Qy        440 ATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATC 499
                |  || ||||| || || |||      ||||||  ||   |||    ||| | |
Db    1393166 CAGAAGGCATGGTAGGTGGGCAACAAGCCGATATGGAAGCTGAAAATAAACAAGTCACAT 1393225

Qy        500 TTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTA 559
              | |||    | | |   || |||       |||||  || | | |||||| ||| || || |
Db    1393226 TAGAAGAACTAGCATCAATCCACGCTCGCAAAACTGGTGAACTATTAATTTTCGCAGTAA 1393285

Qy        560 TGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAGAAAGTTATA 619
                 || ||| | |||||| | |     |      |   | | |||  |||  || || |
Db    1393286 CCTCAGCCGCAAAAATCGCTGAAGCAACTCCAGAACAAACAAAACGATTACGAATTTTTG 1393345

Qy        620 GTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCTATGGTGATG 679
              |  || | || ||    || || ||||  || ||| | |||||      ||| ||||
Db    1393346 CAGAAAATATCGGCATTGGTTTTCAAATTAGCGACGATATTTTAGATGTAATTGGCGATG 1393405

Qy        680 AAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTACGTACGTGA 739
              || ||||  | ||||||||     ||   |||   |       |||||||||| ||
Db    1393406 AAACAAAAATGGGTAAAAAGACAGGGGTCGACGCTTTTCTGAATAAAAGTACCTATCCCG 1393465

Qy        740 GTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTATCAT 783
              | ||| |       || || ||  ||     ||| | | | |||
Db    1393466 GATTACTCACACTCGAAGGTGCTAAACGGGCATTAAATGAGCAT 1393509


Search completed: May 30, 2003, 16:15:17
Job time : 287 secs
```



GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 30, 2003, 16:10:29 ; Search time 61 Seconds
(without alignments)
4328.667 Million cell updates/sec

Title: US-09-925-637-63
Perfect score: 861
Sequence: 1 atgacgaatctaccgatgaa.....tattagaaatcgttgattta 861

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	853	99.1	882	3	US-09-276-873-1	Sequence 1, Appli
2	449.8	52.2	909	4	US-09-134-001C-334	Sequence 334, App
3	157.2	18.3	894	1	US-08-534-910B-5	Sequence 5, Appli
4	155.6	18.1	893	1	US-08-333-321-1	Sequence 1, Appli
5	155.6	18.1	894	1	US-08-534-910B-2	Sequence 2, Appli
6	155.6	18.1	894	1	US-08-534-910B-3	Sequence 3, Appli
7	155.6	18.1	894	1	US-08-534-910B-4	Sequence 4, Appli
8	155.6	18.1	894	3	US-08-886-466-1	Sequence 1, Appli
9	155.6	18.1	894	4	US-09-475-304-1	Sequence 1, Appli
10	155.6	18.1	894	4	US-09-101-126-4	Sequence 4, Appli
11	155.6	18.1	894	4	US-09-367-528A-4	Sequence 4, Appli
12	154	17.9	894	1	US-08-534-910B-1	Sequence 1, Appli
13	154	17.9	894	4	US-09-367-528A-2	Sequence 2, Appli
14	150.6	17.5	486	3	US-09-217-609A-28	Sequence 28, Appl
15	150.6	17.5	486	4	US-08-873-235B-28	Sequence 28, Appl
16	136.4	15.8	885	3	US-09-187-050-11	Sequence 11, Appl
17	134.8	15.7	1179	3	US-09-187-050-13	Sequence 13, Appl
18	134.8	15.7	1179	3	US-09-187-050-15	Sequence 15, Appl
19	134.8	15.7	1179	3	US-09-187-050-17	Sequence 17, Appl
20	134.8	15.7	1179	3	US-09-187-050-19	Sequence 19, Appl
21	134.8	15.7	1179	3	US-09-187-050-21	Sequence 21, Appl
22	134.8	15.7	1179	3	US-09-187-050-23	Sequence 23, Appl
23	134.8	15.7	1179	3	US-09-187-050-25	Sequence 25, Appl
24	134.8	15.7	1889	3	US-09-187-050-1	Sequence 1, Appli
25	124.6	14.5	876	3	US-09-275-742-1	Sequence 1, Appli
c 26	124.2	14.4	10011	4	US-08-961-527-76	Sequence 76, Appl
27	76.8	8.9	1284	4	US-08-858-207A-173	Sequence 173, App
28	72.6	8.4	861	3	US-09-187-050-4	Sequence 4, Appli
29	72	8.4	909	1	US-07-783-705A-7	Sequence 7, Appli
30	72	8.4	6918	1	US-07-783-705A-13	Sequence 13, Appl
31	71.4	8.3	990	1	US-08-410-167A-1	Sequence 1, Appli
32	71.4	8.3	993	1	US-08-705-377-1	Sequence 1, Appli
33	71.4	8.3	993	1	US-08-705-377-2	Sequence 2, Appli
34	71.4	8.3	993	1	US-08-705-377-4	Sequence 4, Appli
35	71.4	8.3	993	1	US-08-705-377-5	Sequence 5, Appli
36	71.4	8.3	993	2	US-09-052-962-1	Sequence 1, Appli
37	71.4	8.3	993	2	US-09-052-962-2	Sequence 2, Appli
38	71.4	8.3	993	2	US-09-052-962-4	Sequence 4, Appli
39	71.4	8.3	993	2	US-09-052-962-5	Sequence 5, Appli
40	71.4	8.3	993	2	US-09-053-068-1	Sequence 1, Appli
41	71.4	8.3	993	2	US-09-053-068-2	Sequence 2, Appli
42	71.4	8.3	993	2	US-09-053-068-4	Sequence 4, Appli
43	71.4	8.3	993	2	US-09-053-068-5	Sequence 5, Appli
44	71.4	8.3	993	2	US-08-898-560-2	Sequence 2, Appli
45	71.4	8.3	993	4	US-09-101-126-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-276-873-1
; Sequence 1, Application US/09276873
; Patent No. 6107058
; GENERAL INFORMATION:
; APPLICANT: Wilding, Edwina Imogen
; APPLICANT: Gwynn, Michael
; TITLE OF INVENTION: IspA
; FILE REFERENCE: GM10208
; CURRENT APPLICATION NUMBER: US/09/276,873
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 882
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-276-873-1

Query Match 99.1%; Score 853; DB 3; Length 882;
Best Local Similarity 99.4%; Pred. No. 2.6e-198;
Matches 856; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGCATGTTGTATTCATTAAATGCT 120

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTGTGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              |||||||||||||||||||||||||||||||||||| | ||||||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAAGGAAATTAACAAATCAT 300

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
```

```
Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAA 600

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

          721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||||||||||||||||||||
Db        841 TTATTAGAAATCGTTGATTTA 861
```

RESULT 2
US-09-134-001C-334
; Sequence 334, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 334
; LENGTH: 909
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-334

Query Match 52.2%; Score 449.8; DB 4; Length 909;
Best Local Similarity 70.2%; Pred. No. 1.6e-100;
Matches 604; Conservative 0; Mismatches 257; Indels 0; Gaps 0;

```
Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              |||| ||| |||| ||||||||||||||||| ||  | | ||||    | |        |
Db         28 ATGAAGAAACTACAGATGAATAAATTAATAAATATAATTAATACATCACTGAATAAGTCA 87

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||| |   ||||  | |  | ||| |  ||||||||||||||   ||||||||||||||
Db         88 ATACAATCATCACCATTAAAAACTAATTTAGAAGAAAGTATGAAATATTCATTAAATGCT 147

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              || |||||| | ||| |||||||  | ||||||| | ||  || |    || || |  ||
Db        148 GGTGGTAAAAGAATCAGACCAGTCATATTATTATTAACACTAAAAATGCTTAACAAAGAT 207

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              ||| |   |||  | || || ||  | ||| | ||||||||||||||| |||||  | |||
Db        208 TATCAACAAGGACTAAATAGTGCTTTAGCATTGGAAATGATTCATACTTATTCTTTAATT 267

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              ||||||||  |||||| ||||||||||| ||||| ||  |||||||||||||||||||||
Db        268 CATGATGATTTACCAGCAATGGATAATGACGATTACCGTAGAGGAAAATTAACAAATCAT 327

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||| |||||||| |||| || ||  | || |||||||| ||||||| |||||  |||
Db        328 AAAGTTTATGGTGAATGGAAAGCCATTCTTGCTGGTGATGCATTATTAACAAAAGCTTTT 387

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||| |  |||| | |||| ||    |   |||   || ||| |||   || || || |
Db        388 GAATTAGTTTCTAATGATACTACCATTGAAGATAGTGTGAAAGTAAGTATTATAAAAAGA 447

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              || |||| ||||||||| ||| | |||||||| || || |||  | | |||||| ||||
Db        448 CTTTCAAAAGCAAGTGGACATTTGGGAATGGTGGGTGGCCAAGCGCTTGATATGGAAAGT 507

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              |||||  |  |||||  | | |||||||| |||   || ||  |||| || ||||| ||
Db        508 GAAGGGAAGTCAATTCGTTTAGAAACTTTAGAATCAATTCATGAAACTAAGACAGGCGCT 567

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              ||| ||| |||| | ||||||  ||| | ||| || ||| | || ||  | |  | |
Db        568 TTACTAAATTTTTCAGTTATGGCTGCGGTAGACATTGCTCAAGTAGAACAAAATATTGCT 627

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              |  |||||||     | |||| |||||||||| |||||||| || ||||||||||||||
Db        628 AAGAATTTAGATGAATTTAGTCATCATTTAGGAATGATGTTTCAAATTAAAGATGATTTA 687

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              | ||    |||||||||||| |||  | || |||||||| |||||  ||| | | ||||
Db        688 CTGGATGTGTATGGTGATGAATCAAAACTTGGCAAAAAAGTAGGCAGTGATATAGTAAAT 747

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              |||||||||| || ||   |||| | || ||||| || |||||||| || || |  ||
Db        748 CATAAAAGTACTTATGTTTCTTTACTTGGAAAAGAAGGAGCAGAAGAAAAGTTAAACAAT 807

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              |||  | |    ||  || |    |||  ||||||  ||| ||||   ||||      |
Db        808 CATCAATATCTTGCTATGAACTGCTTAAATCAAATTTCTGATCAATATGATACTTCTGAA 867

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||   || || || ||||||
Db        868 TTAAGTGATATTGTAGATTTA 888

RESULT 3
US-08-534-910B-5
; Sequence 5, Application US/08534910B
; Patent No. 5766911
; GENERAL INFORMATION:
; APPLICANT: KOIKE, Ayumi
; APPLICANT: OBATA, Shusei
; APPLICANT: NISHINO, Tokuzo
; APPLICANT: OHNUMA, Shinichi
; APPLICANT: NAKAZAWA, Takeshi
; APPLICANT: OGURA, Kyozo
; APPLICANT: KOYAMA, Tanetoshi
; TITLE OF INVENTION: Mutated Farnesyldiphosphate Synthase Capable
; TITLE OF INVENTION: Of Synthesizing Geranylgeranyldiphosphate And Gene Coding T
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1025 Connecticut Avenue, N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.
; ZIP: 20036-5405
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.25" Floppy Disk
; COMPUTER: IBM PC Compatible
```

```
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;      SOFTWARE:  IBM/Word Perfect 6.1 Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/534,910B
;      FILING DATE:  28-SEPT-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-25253
;      FILING DATE:  14-FEB-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Toffenetti, Judith L.
;      REGISTRATION NUMBER:  39,048
;      REFERENCE/DOCKET NUMBER:  77670/398
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)429-1776
;      TELEFAX:  (202)429-0796
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  894 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Bacillus stearothermophilus
US-08-534-910B-5

  Query Match            18.3%;  Score 157.2;  DB 1;  Length 894;
  Best Local Similarity   51.0%;  Pred. No. 1.5e-29;
  Matches  400;  Conservative    0;  Mismatches  378;  Indels    6;  Gaps    1;

Qy         17 TGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTAA 76
              | ||  |   || | | |  ||  | |  |   | ||   |   ||| |   | || |
Db         29 TCAACGAGCAAAAACAGGCGGTGGAAACAGCGCTCTCCCGTTATATAGAGCGCTTAGAAG 88

Qy         77 TGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCC 136
              |    |  ||||  || |    |||  ||| |||||  | || || || ||||| ||||
Db         89 GGCCGGCGAAGCTGAAAAAGGCGATGGCGTACTCATTGGAGGCCGGCGGCAAACGAATCC 148

Qy        137 GACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGA 196
              | ||  | ||| |  |   |||  |  |  | ||    |  ||    | | | ||  ||
Db        149 GTCCGTTGCTGCTTCTGTCCACCGTTCAGGCGCTCGGCAAAGACCCGGCGGTCGGATTGC 208

Qy        197 AGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAG 256
                |||    ||  | ||||||||  ||||| || ||  | || ||||||||  | ||
Db        209 CCGTCGCCTGCGCGATTGAAATGATCCATACGTACTCTTTGATCCATGATGATTTGCCGA 268

Qy        257 CGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGT 316
              ||||| || |||||||  || || || ||    || || ||||||| |  || |||
Db        269 GCATGGACAACGATGATTTGCGGCGCGGCAAGCCGACGAACCATAAAGTGTTCGGCGAGG 328

Qy        317 GGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTC----- 371
              ||  || || || || || || || |  || || ||  | || ||| || | ||  |
Db        329 CGATGGCCATCTTGGCGGGGGACGGGTTGTTGACGTACGCGTTTCAATTGATCACCGAAA 388

Qy        372 -AAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAG 430
               |||||  |  |  ||  |   ||     |     |  |  |||||||| | | ||
Db        389 TCGACGATGAGCGCATCCCTCCTTCCGTCCGGCTTCGGCTCATCGAACGGCTGGCGAAAG 448

Qy        431 CAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAAC 490
              |   ||||  |  || |||||||  ||||||       ||||||  || | || ||  ||
Db        449 CGGCCGGTCCGGAAGGGATGGTCGCCGGTCAGGCAGCCGATATGGAAGGAGAGGGGAAAA 508

Qy        491 CAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTT 550
              |  |   |||       | |||   || ||       ||||| ||   |' |  |   |
Db        509 CGCTGACGCTTTCGGAGCTCGAATACATTCATCGGCATAAAACCGGGAAAATGCTGCAAT 568

Qy        551 TTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAG 610
                ||       |  ||   ||||    |  ||||        |      |  | |
Db        569 ACAGCGTGCACGCCGGCGCCTTGATCGGCGGCGCTGATGCCCGGCAAACGCGGGAGCTTG 628

Qy        611 AAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCT 670
              |   |        ||| |||| |   || || |||   |||||| | . | ||
Db        629 ACGAATTCGCCGCCCATCTAGGCCTTGCCTTTCAAATTCGCGATGATATTCTCGATATTG 688

Qy        671 ATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTA 730
              | || |  |||| |||  | || ||    || |||||||| |     || || |||  |
Db        689 AAGGGGCAGAAGAAAAAATCGGCAAGCCGGTCGGCAGCGACCAAAGCAACAACAAAGCGA 748

Qy        731 CGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGACG 790
              ||||       ||  |   |   |  |||||  | || || ||| | |  ||||  || |
Db        749 CGTATCCAGCGTTGCTGTCGCTTGCCGGCGCGAAGGAAAAGTTGGCGTTCCATATCGAGG 808

Qy        791 CAGC 794
              | ||
Db        809 CGGC 812


RESULT 4
US-08-333-321-1
; Sequence 1, Application US/08333321
; Patent No. 5786192
;  GENERAL INFORMATION:
;    APPLICANT:  Obata, Shusei
;    TITLE OF INVENTION:  FARNESYL PYROPHOPHATE SYNTHETASE
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Edward W. Greason, Esq.
;      STREET:  1 Broadway
;      CITY:  New York
;      STATE:  N.Y.
;      COUNTRY:  U.S.A.
;      ZIP:  10004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  No. 5786192epad, Windows 3.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/333,321
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/953,424
;      FILING DATE:  29-SEP-1992
;      APPLICATION NUMBER:  JP 3-253788
;      FILING DATE:  01-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Greason, Edward W.
;      REGISTRATION NUMBER:  18,918
;      REFERENCE/DOCKET NUMBER:  077670/00310
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 425-7200
;      TELEFAX:  (212) 425-5288
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  893 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  Bacillus stearothermophilus
US-08-333-321-1

  Query Match            18.1%;  Score 155.6;  DB 1;  Length 893;
  Best Local Similarity   50.9%;  Pred. No. 3.7e-29;
  Matches  399;  Conservative    0;  Mismatches  379;  Indels    6;  Gaps    1;

Qy         17 TGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTAA 76
              | ||  |   || | | |  ||  | |  |   | ||   |   ||| |   | || |
Db         29 TCAACGAGCAAAAACAGGCGGTGGAAACAGCGCTCTCCCGTTATATAGAGCGCTTAGAAG 88

Qy         77 TGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCC 136
              |    |  ||||  || |    |||  ||| |||||  | || || || ||||| ||||
Db         89 GGCCGGCGAAGCTGAAAAAGGCGATGGCGTACTCATTGGAGGCCGGCGGCAAACGAATCC 148
```

```
Qy        137 GACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGA 196
              | ||  | ||| |  |   |||  |     | ||     |  ||    | | | ||  ||
Db        149 GTCCGTTGCTGCTTCTGTCCACCGTTCGGGCGCTCGGCAAAGACCCGGCGGTCGGATTGC 208

Qy        197 AGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAG 256
                 |||    ||  | ||||||||| ||||| || ||  | || ||||||||  | ||
Db        209 CCGTCGCCTGCGCGATTGAAATGATCCATACGTACTCTTTGATCCATGATGATTTGCCGA 268

Qy        257 CGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGT 316
               ||||| || |||||||  || || || ||    || || ||||||||  |  || |||
Db        269 GCATGGACAACGATGATTTGCGGCGCGGCAAGCCGACGAACCATAAAGTGTTCGGCGAGG 328

Qy        317 GGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTC----- 371
               ||  || || ||  || || ||  |  || || ||  | || ||| || | ||  |
Db        329 CGATGGCCATCTTGGCGGGGGACGGGTTGTTGACGTACGCGTTTCAATTGATCACCGAAA 388

Qy        372 -AAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAG 430
                |||||  |  |  ||  |   ||     |  |  |||||||| | | ||
Db        389 TCGACGATGAGCGCATCCCTCCTTCCGTCCGGCTTCGGCTCATCGAACGGCTGGCGAAAG 448

Qy        431 CAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAAC 490
              |    ||||  |  || |||||||  ||||||       ||||||  || | ||  ||  ||
Db        449 CGGCCGGTCCGGAAGGGATGGTCGCCGGTCAGGCAGCCGATATGGAAGGAGAGGGGAAAA 508

Qy        491 CAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTT 550
              | |   |||       | |||   || ||       |||||  ||   |  |  |    |
Db        509 CGCTGACGCTTTCGGAGCTCGAATACATTCATCGGCATAAAACCGGGAAAATGCTGCAAT 568

Qy        551 TTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAG 610
                 ||     |  ||    ||||    |  ||||        |     |  | | |
Db        569 ACAGCGTGCACGCCGGCGCCTTGATCGGCGGCGCTGATGCCCGGCAAACGCGGGAGCTTG 628

Qy        611 AAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCT 670
              |   |        ||| ||||  |    || || |||   ||||||  |  | ||
Db        629 ACGAATTCGCCGCCCATCTAGGCCTTGCCTTTCAAATTCGCGATGATATTCTCGATATTG 688

Qy        671 ATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTA 730
              | || |  |||| |||  | || ||     || |||||||| |     || || |||   |
Db        689 AAGGGGCAGAAGAAAAAATCGGCAAGCCGGTCGGCAGCGACCAAAGCAACAACAAAGCGA 748

Qy        731 CGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGACG 790
              ||||       ||  |   |  |  |||||  | || || ||| | |  |||  || |
Db        749 CGTATCCAGCGTTGCTGTCGCTTGCCGGCGCGAAGGAAAAGTTGGCGTTCCATATCGAGG 808

Qy        791 CAGC 794
              | ||
          809 CGGC 812


RESULT 5
US-08-534-910B-2
; Sequence 2, Application US/08534910B
; Patent No. 5766911
;  GENERAL INFORMATION:
;    APPLICANT:  KOIKE, Ayumi
;    APPLICANT:  OBATA, Shusei
;    APPLICANT:  NISHINO, Tokuzo
;    APPLICANT:  OHNUMA, Shinichi
;    APPLICANT:  NAKAZAWA, Takeshi
;    APPLICANT:  OGURA, Kyozo
;    APPLICANT:  KOYAMA, Tanetoshi
;    TITLE OF INVENTION:  Mutated Farnesyldiphosphate Synthase Capable
;    TITLE OF INVENTION:  Of Synthesizing Geranylgeranyldiphosphate And Gene Coding Ther
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenyon & Kenyon
;      STREET:  1025 Connecticut Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  U.S.
;      ZIP:  20036-5405
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.25" Floppy Disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;      SOFTWARE:  IBM/Word Perfect 6.1 Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/534,910B
;      FILING DATE:  28-SEPT-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-25253
;      FILING DATE:  14-FEB-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Toffenetti, Judith L.
;      REGISTRATION NUMBER:  39,048
;      REFERENCE/DOCKET NUMBER:  77670/398
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)429-1776
;      TELEFAX:  (202)429-0796
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  894 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Bacillus stearothermophilus
US-08-534-910B-2

  Query Match            18.1%;  Score 155.6;  DB 1;  Length 894;
  Best Local Similarity  50.9%;  Pred. No. 3.7e-29;
  Matches  399;  Conservative    0;  Mismatches  379;  Indels    6;  Gaps    1;

Qy         17 TGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTAA 76
              | ||  |   || | | |  || | |  |   | ||   |   ||| |    | || |
Db         29 TCAACGAGCAAAAACAGGCGGTGGAAACAGCGCTCTCCCGTTATATAGAGCGCTTAGAAG 88

Qy         77 TGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCC 136
                |    |  || |  || |    |||  ||| |||||  | || || || ||||| ||||
Db         89 GGCCGGCGAAGGTGAAAAAGGCGATGGCGTACTCATTGGAGGCCGGCGGCAAACGAATCC 148

Qy        137 GACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGA 196
              | ||  | ||| |  |   |||  |     | ||     |  ||    | | | ||  ||
Db        149 GTCCGTTGCTGCTTCTGTCCACCGTTCAGGCGCTCGGCAAAGACCCGGCGGTCGGATTGC 208

Qy        197 AGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAG 256
                 |||    ||  | ||||||||| ||||| || ||  | || ||||||||  | ||
Db        209 CCGTCGCCTGCGCGATTGAAATGATCCATACGTACTCTTTGATCCATGATGATTTGCCGA 268

Qy        257 CGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGT 316
               ||||| || |||||||  || || || ||    || || ||||||||  |  || |||
Db        269 GCATGGACAACGATGATTTGCGGCGCGGCAAGCCGACGAACCATAAAGTGTTCGGCGAGG 328

Qy        317 GGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTC----- 371
               ||  || || ||  || || ||  |  || || ||  | || ||| || | ||  |
Db        329 CGATGGCCATCTTGGCGGGGGACGGGTTGTTGACGTACGCGTTTCAATTGATCACCGAAA 388

Qy        372 -AAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAG 430
                |||||  |  |  ||  |   ||     |  |  |||||||| | | ||
Db        389 TCGACGATGAGCGCATCCCTCCTTCCGTCCGGCTTCGGCTCATCGAACGGCTGGCGAAAG 448

Qy        431 CAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAAC 490
              |    ||||  |  || |||||||  ||||||       ||||||  || | ||  ||  ||
Db        449 CGGCCGGTCCGGAAGGGATGGTCGCCGGTCAGGCAGCCGATATGGAAGGAGAGGGGAAAA 508

Qy        491 CAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTT 550
              | |   |||       | |||   || ||       |||||  ||   |  |  |    |
Db        509 CGCTGACGCTTTCGGAGCTCGAATACATTCATCGGCATAAAACCGGGAAAATGCTGCAAT 568

Qy        551 TTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAG 610
                 ||     |  ||    ||||    |  ||||        |     |  | | |
Db        569 ACAGCGTGCACGCCGGCGCCTTGATCGGCGGCGCTGATGCCCGGCAAACGCGGGAGCTTG 628
```

```
Qy        611 AAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCT 670
              |    |       ||| |||| |    || || |||   |||||| |  | ||
Db        629 ACGAATTCGCCGCCCATCTAGGCCTTGCCTTTCAAATTCGCGATGATATTCTCGATATTG 688

Qy        671 ATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTA 730
              | || |  |||| |||  | || ||    || ||||||||  |     || || |||   |
Db        689 AAGGGGCAGAAGAAAAAATCGGCAAGCCGGTCGGCAGCGACCAAAGCAACAACAAAGCGA 748

Qy        731 CGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGACG 790
              ||||       ||  |   |  ||||| | || || ||| | |  ||||  || |
Db        749 CGTATCCAGCGTTGCTGTCGCTTGCCGGCGCGAAGGAAAAGTTGGCGTTCCATATCGAGG 808

Qy        791 CAGC 794
              | ||
Db        809 CGGC 812


RESULT 6
US-08-534-910B-3
; Sequence 3, Application US/08534910B
; Patent No. 5766911
; GENERAL INFORMATION:
;   APPLICANT:  KOIKE, Ayumi
;   APPLICANT:  OBATA, Shusei
;   APPLICANT:  NISHINO, Tokuzo
;   APPLICANT:  OHNUMA, Shinichi
;   APPLICANT:  NAKAZAWA, Takeshi
;   APPLICANT:  OGURA, Kyozo
;   APPLICANT:  KOYAMA, Tanetoshi
;   TITLE OF INVENTION:  Mutated Farnesyldiphosphate Synthase Capable
;   TITLE OF INVENTION:  Of Synthesizing Geranylgeranyldiphosphate And Gene Coding Ther
;   NUMBER OF SEQUENCES:  10
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Kenyon & Kenyon
;     STREET:  1025 Connecticut Avenue, N.W., Suite 600
;     CITY:  Washington
;     STATE:  DC
;     COUNTRY:  U.S.
;     ZIP:  20036-5405
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  3.25" Floppy Disk
;     COMPUTER:  IBM PC Compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;     SOFTWARE:  IBM/Word Perfect 6.1 Windows
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/534,910B
;     FILING DATE:  28-SEPT-1995
;     CLASSIFICATION:  435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  JP 7-25253
;     FILING DATE:  14-FEB-1995
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Toffenetti, Judith L.
;     REGISTRATION NUMBER:  39,048
;     REFERENCE/DOCKET NUMBER:  77670/398
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (202)429-1776
;     TELEFAX:  (202)429-0796
; INFORMATION FOR SEQ ID NO:  3:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  894 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  double
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  DNA
;   ORIGINAL SOURCE:
;     ORGANISM:  Bacillus stearothermophilus
US-08-534-910B-3

  Query Match           18.1%;  Score 155.6;  DB 1;  Length 894;
  Best Local Similarity  50.9%;  Pred. No. 3.7e-29;
  Matches  399;  Conservative    0;  Mismatches  379;  Indels    6;  Gaps    1;

Qy         17 TGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTAA 76
              | ||  |   || | | |  ||  | |  |   | ||   |   ||| |    | || |
Db         29 TCAACGAGCAAAAACAGGCGGTGGAAACAGCGCTCTCCCGTTATATAGAGCGCTTAGAAG 88

Qy         77 TGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCC 136
              |    | ||||  || |    |||  ||| |||||  | || || || ||||| ||||
Db         89 GGCCGGCGAAGCTGAAAAAGGCGATGGCGTACTCATTGGAGGCCGGCGGCAAACGAATCC 148

Qy        137 GACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGA 196
              | ||  | ||| |  |   |||  |     | ||    |  ||    | | | ||  ||
Db        149 GTCCGTTGCTGCTTCTGTCCACCGTTCGGGCGCTCGGAAAAGACCCGGCGGTCGGATTGC 208

Qy        197 AGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAG 256
                 |||    ||  | ||||||| |||| || ||  | || ||||||||  | ||
Db        209 CCGTCGCCTGCGCGATTGAAATGATCCATACGTACTCTTTGATCCATGATGATTTGCCGA 268

Qy        257 CGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGT 316
              |||||  || |||||||  || || || ||    || || ||||||||  |  || |||
Db        269 GCATGGACAACGATGATTTGCGGCGCGGCAAGCCGACGAACCATAAAGTGTTCGGCGAGG 328

Qy        317 GGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTC----- 371
              ||  || || || || || || |  || || ||  | || ||| || | | ||  |
Db        329 CGATGGCCATCTTGGCGGGGGACGGGTTGTTGACGTACGCGTTTCAATTGATCACCGAAA 388

Qy        372 -AAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAG 430
               |||||  |  |  ||    |     |  |  |||||||| | | ||
Db        389 TCGACGATGAGCGCATCCCTCCTTCCGTCCGGCTTCGGCTCATCGAACGGCTGGCGAAAG 448

Qy        431 CAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAAC 490
              |    ||||  |  || |||| || ||||||        ||||||  || | || ||  ||
Db        449 CGGCCGGTCCGGAAGGGATGGCCGCCGGTCAGGCAGCCGATATGGAAGGAGAGGGGAAAA 508

Qy        491 CAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTT 550
              |  |   |||       | |||   || ||       ||||| ||  | | |    |
Db        509 CGCTGACGCTTTCGGAGCTCGAATACATTCATCGGTATAAAACCGGGAAAATGCTGCAAT 568

Qy        551 TTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAG 610
                 ||      |  ||    ||||    |  ||||         |      |  | |
Db        569 ACAGCGTGCACGCCGGCGCCTTGATCGGCGGCGCTGATGCCCGGCAAACGCGGGAGCTTG 628

Qy        611 AAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCT 670
              |    |         ||| ||||  |    || || |||   ||||||  |  | ||
Db        629 ACGAATTCGCCGCCCATCTAGGCCTTGCCTTTCAAATTCGCGATGATATTCTCGATATTG 688

Qy        671 ATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTA 730
              | || |  |||| |||  | || ||    || ||||||||  |     || || |||   |
Db        689 AAGGGGCAGAAGAAAAAATCGGCAAGCCGGTCGGCAGCGACCAAAGCAACAACAAAGCGA 748

Qy        731 CGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGACG 790
              ||||       ||  |   |  |||||| | || || ||| | |  ||||  || |
Db        749 CGTATCCAGCGTTGCTGTCGCTTGCCGGCGCAAAGGAAAAGTTGGCGTTCCATATCGAGG 808

Qy        791 CAGC 794
              | ||
Db        809 CGGC 812


RESULT 7
US-08-534-910B-4
; Sequence 4, Application US/08534910B
; Patent No. 5766911
; GENERAL INFORMATION:
;   APPLICANT:  KOIKE, Ayumi
;   APPLICANT:  OBATA, Shusei
;   APPLICANT:  NISHINO, Tokuzo
;   APPLICANT:  OHNUMA, Shinichi
;   APPLICANT:  NAKAZAWA, Takeshi
;   APPLICANT:  OGURA, Kyozo
;   APPLICANT:  KOYAMA, Tanetoshi
;   TITLE OF INVENTION:  Mutated Farnesyldiphosphate Synthase Capable
;   TITLE OF INVENTION:  Of Synthesizing Geranylgeranyldiphosphate And Gene Coding T
;   NUMBER OF SEQUENCES:  10
;   CORRESPONDENCE ADDRESS:
```

```
;       ADDRESSEE:  Kenyon & Kenyon
;       STREET:  1025 Connecticut Avenue, N.W., Suite 600
;       CITY:  Washington
;       STATE:  DC
;       COUNTRY:  U.S.
;       ZIP:  20036-5405
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.25" Floppy Disk
;       COMPUTER:  IBM PC Compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;       SOFTWARE:  IBM/Word Perfect 6.1 Windows
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/534,910B
;       FILING DATE:  28-SEPT-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  JP 7-25253
;       FILING DATE:  14-FEB-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Toffenetti, Judith L.
;       REGISTRATION NUMBER:  39,048
;       REFERENCE/DOCKET NUMBER:  77670/398
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202)429-1776
;       TELEFAX:  (202)429-0796
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  894 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
;     ORIGINAL SOURCE:
;       ORGANISM:  Bacillus stearothermophilus
US-08-534-910B-4

  Query Match             18.1%;  Score 155.6;  DB 1;  Length 894;
  Best Local Similarity   50.9%;  Pred. No. 3.7e-29;
  Matches  399;  Conservative    0;  Mismatches  379;  Indels    6;  Gaps    1;

Qy         17 TGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTAA 76
Db         29 TCAACGAGCAAAAACAGGCGGTGGAAACAGCGCTCTCCCGTTATATAGAGCGCTTAGAAG 88

Qy         77 TGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCC 136
Db         89 GGCCGGCGAAGCTGAAAAAGGCGATGGCGTACTCATTGGAGGCCGGCGGCAAACGAATCC 148

Qy        137 GACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGA 196
Db        149 GTCCGTTGCTGCTTCTGTCCACCGTTCGGGCGCTCGGCAAAGACCCGGCGGTCGGATTGC 208

Qy        197 AGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAG 256
Db        209 CCGTCGCCTGCGCGATTGAAATGATCCATACGCACTCTTTGATCCATGATGATTTGCCGA 268

Qy        257 CGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGT 316
Db        269 GCATGGACAACGATGATTTGCGGCGCGGCAAGCCGACGAACCATAAAGTGTTCGGCGAGG 328

Qy        317 GGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTC----- 371
Db        329 CGATGGCCATCTTGGCGGGGGACGGGTTGTTGACGTACGCGTTTCAATTGATCACCGAAA 388

Qy        372 -AAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAG 430
Db        389 TCGACGATGAGCGCATCCCTCCTTCCGTCCGGCTTCGGCTCATCGAACGGCTGGCGAAAG 448

Qy        431 CAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAAC 490
Db        449 CGGCCGGTCCGGAAGGGATGGTCGCCGGTCAGGCAGCCGATATGGAAGGAGAGGGGAAAA 508

Qy        491 CAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTT 550
Db        509 CGCTGACGCTTTCGGAGCTCGAATACATTCATCGGCATAAAACCGGGAAAATGCTGCAAT 568

Qy        551 TTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAG 610
Db        569 ACAGCGTGCACGCCGGCGCCTTGATCGGCGGCGCTGATGCCCGGCAAACGCGGGAGCTTG 628

Qy        611 AAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCT 670
Db        629 ACGAATTCGCCGCCCATCTAGGCCTTGCCTTTCAAATTCGCGATGATATTCTCGATATTG 688

Qy        671 ATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTA 730
Db        689 AAGGGGCAGAAGAAAAAATCGGCAAGCGGGTCGGCAGCGACCAAAGCAACAACAAAGCGA 748

Qy        731 CGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGACG 790
Db        749 CGTATCCAGCGTTGCTGTCGCTTGCCGGCGCGAAGGAAAAGTTGACGTTCCATATCGAGG 808

Qy        791 CAGC 794
Db        809 CGGC 812


RESULT 8
US-08-886-466-1
; Sequence 1, Application US/08886466C
; Patent No. 6040165
; GENERAL INFORMATION:
;  APPLICANT: Narita, Keishi
;  APPLICANT: Ishida, Chika
;  APPLICANT: Takeuchi, Yoshie
;  APPLICANT: Ohto, Chikara
;  APPLICANT: Ohnuma, Shinichi
;  APPLICANT: Nishino, Tokuzo
;  TITLE OF INVENTION: MUTANT PRENYL DIPHOSPHATE SYNTHASE
;  FILE REFERENCE: 77670/494
;  CURRENT APPLICATION NUMBER: US/08/886,466C
;  CURRENT FILING DATE: 1997-07-10
;  EARLIER APPLICATION NUMBER: JP 8-191635
;  EARLIER FILING DATE: 1996-07-03
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 894
;   TYPE: DNA
;   ORGANISM: Bacillus stearothermophilus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(891)
US-08-886-466-1

  Query Match             18.1%;  Score 155.6;  DB 3;  Length 894;
  Best Local Similarity   50.9%;  Pred. No. 3.7e-29;
  Matches  399;  Conservative    0;  Mismatches  379;  Indels    6;  Gaps    1;

Qy         17 TGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTAA 76
Db         29 TCAACGAGCAAAAACAGGCGGTGGAAACAGCGCTCTCCCGTTATATAGAGCGCTTAGAAG 88

Qy         77 TGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCC 136
Db         89 GGCCGGCGAAGCTGAAAAAGGCGATGGCGTACTCATTGGAGGCCGGCGGCAAACGAATCC 148

Qy        137 GACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGA 196
Db        149 GTCCGTTGCTGCTTCTGTCCACCGTTCGGGCGCTCGGCAAAGACCCGGCGGTCGGATTGC 208

Qy        197 AGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAG 256
Db        209 CCGTCGCCTGCGCGATTGAAATGATCCATACGTACTCTTTGATCCATGATGATTTGCCGA 268

Qy        257 CGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGT 316
```

```
Db        269 GCATGGACAACGATGATTTGCGGCGCGGCAAGCCGACGAACCATAAAGTGTTCGGCGAGG 328

Qy        317 GGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTC----- 371

Db        329 CGATGGCCATCTTGGCGGGGGACGGGTTGTTGACGTACGCGTTTCAATTGATCACCGAAA 388

Qy        372 -AAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAG 430

Db        389 TCGACGATGAGCGCATCCCTCCTTCCGTCCGGCTTCGGCTCATCGAACGGCTGGCGAAAG 448

Qy        431 CAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAAC 490

Db        449 CGGCCGGTCCGGAAGGGATGGTCGCCGGTCAGGCAGCCGATATGGAAGGAGAGGGGAAAA 508

Qy        491 CAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTT 550

Db        509 CGCTGACGCTTTCGGAGCTCGAATACATTCATCGGCATAAAACCGGGAAAATGCTGCAAT 568

Qy        551 TTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAG 610

Db        569 ACAGCGTGCACGCCGGCGCCTTGATCGGCGGCGCTGATGCCCGGCAAACGCGGGAGCTTG 628

Qy        611 AAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCT 670

Db        629 ACGAATTCGCCGCCCATCTAGGCCTTGCCTTTCAAATTCGCGATGATATTCTCGATATTG 688

Qy        671 ATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTA 730

Db        689 AAGGGGCAGAAGAAAAAATCGGCAAGCCGGTCGGCAGCGACCAAAGCAACAACAAAGCGA 748

Qy        731 CGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGACG 790

Db        749 CGTATCCAGCGTTGCTGTCGCTTGCCGGCGCGAAGGAAAAGTTGGCGTTCCATATCGAGG 808

Qy        791 CAGC 794

Db        809 CGGC 812


RESULT 9
US-09-475-304-1
; Sequence 1, Application US/09475304
; Patent No. 6225096
; GENERAL INFORMATION:
;  APPLICANT: Narita, Keishi
;  APPLICANT: Ishida, Chika
;  APPLICANT: Takeuchi, Yoshie
;  APPLICANT: Ohto, Chikara
;  APPLICANT: Ohnuma, Shinichi
;  APPLICANT: Nishino, Tokuzo
;  TITLE OF INVENTION: MUTANT PRENYL DIPHOSPHATE SYNTHASE
;  FILE REFERENCE: 77670/494
;  CURRENT APPLICATION NUMBER: US/09/475,304
;  CURRENT FILING DATE: 1999-12-30
;  EARLIER APPLICATION NUMBER: JP 8-191635
;  EARLIER FILING DATE: 1996-07-03
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 894
;   TYPE: DNA
;   ORGANISM: Bacillus stearothermophilus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(891)
US-09-475-304-1

  Query Match           18.1%;  Score 155.6;  DB 4;  Length 894;
  Best Local Similarity  50.9%;  Pred. No. 3.7e-29;
  Matches  399;  Conservative    0;  Mismatches  379;  Indels    6;  Gaps    1;

Qy         17 TGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTAA 76

Db         29 TCAACGAGCAAAAACAGGCGGTGGAAACAGCGCTCTCCCGTTATATAGAGCGCTTAGAAG 88

Qy         77 TGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCC 136

Db         89 GGCCGGCGAAGCTGAAAAAGGCGATGGCGTACTCATTGGAGGCCGGCGGCAAACGAATCC 148

Qy        137 GACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGA 196

Db        149 GTCCGTTGCTGCTTCTGTCCACCGTTCGGGCGCTCGGCAAAGACCCGGCGGTCGGATTGC 208

Qy        197 AGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAG 256

Db        209 CCGTCGCCTGCGCGATTGAAATGATCCATACGTACTCTTTGATCCATGATGATTTGCCGA 268

Qy        257 CGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGT 316

Db        269 GCATGGACAACGATGATTTGCGGCGCGGCAAGCCGACGAACCATAAAGTGTTCGGCGAGG 328

Qy        317 GGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTC----- 371

Db        329 CGATGGCCATCTTGGCGGGGGACGGGTTGTTGACGTACGCGTTTCAATTGATCACCGAAA 388

Qy        372 -AAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAG 430

Db        389 TCGACGATGAGCGCATCCCTCCTTCCGTCCGGCTTCGGCTCATCGAACGGCTGGCGAAAG 448

Qy        431 CAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAAC 490

Db        449 CGGCCGGTCCGGAAGGGATGGTCGCCGGTCAGGCAGCCGATATGGAAGGAGAGGGGAAAA 508

Qy        491 CAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTT 550

Db        509 CGCTGACGCTTTCGGAGCTCGAATACATTCATCGGCATAAAACCGGGAAAATGCTGCAAT 568

Qy        551 TTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAG 610

Db        569 ACAGCGTGCACGCCGGCGCCTTGATCGGCGGCGCTGATGCCCGGCAAACGCGGGAGCTTG 628

Qy        611 AAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCT 670

Db        629 ACGAATTCGCCGCCCATCTAGGCCTTGCCTTTCAAATTCGCGATGATATTCTCGATATTG 688

Qy        671 ATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTA 730

Db        689 AAGGGGCAGAAGAAAAAATCGGCAAGCCGGTCGGCAGCGACCAAAGCAACAACAAAGCGA 748

Qy        731 CGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGACG 790

Db        749 CGTATCCAGCGTTGCTGTCGCTTGCCGGCGCGAAGGAAAAGTTGGCGTTCCATATCGAGG 808

Qy        791 CAGC 794

Db        809 CGGC 812


RESULT 10
US-09-101-126-4
; Sequence 4, Application US/09101126
; Patent No. 6316216
; GENERAL INFORMATION:
;  APPLICANT: OHTO, CHIKARA
;  APPLICANT: NAKANE, HIROYUKI
;  APPLICANT: NISHINO, TOKUZO
;  APPLICANT: OHNUMA, SHINICHI
;  APPLICANT: HIROOKA, KAZUTAKE
;  TITLE OF INVENTION: MUTATED PRENYL DIPHOSPHATE SYNTHASES
;  FILE REFERENCE: 77670/566
;  CURRENT APPLICATION NUMBER: US/09/101,126
;  CURRENT FILING DATE: 1999-04-27
;  EARLIER APPLICATION NUMBER: PCT/JP97/03921
;  EARLIER FILING DATE: 1997-10-29
;  EARLIER APPLICATION NUMBER: JP 8-307506
```

```
;  EARLIER FILING DATE: 1996-11-05
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 894
;   TYPE: DNA
;   ORGANISM: Bacillus stearothermophilus
;   FEATURE:
;   OTHER INFORMATION: 256-276 is an Asp-rich coding domain
US-09-101-126-4

  Query Match             18.1%;  Score 155.6;  DB 4;  Length 894;
  Best Local Similarity   50.9%;  Pred. No. 3.7e-29;
  Matches  399;  Conservative    0;  Mismatches  379;  Indels    6;  Gaps    1;

Qy         17 TGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTAA 76
              | ||  |   || | |  ||  | |  |   | ||   |   ||| |    | || |
Db         29 TCAACGAGCAAAAACAGGCGGTGGAAACAGCGCTCTCCCGTTATATAGAGCGCTTAGAAG 88

Qy         77 TGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCC 136
              |    |  ||||  || |    |||  ||| ||||  | || || || ||||| ||||
Db         89 GGCCGGCGAAGCTGAAAAAGGCGATGGCGTACTCATTGGAGGCCGGCGGCAAACGAATCC 148

Qy        137 GACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGA 196
              | ||  | ||| |  |   |||  |     | ||    |  ||    | | | ||  ||
Db        149 GTCCGTTGCTGCTTCTGTCCACCGTTCGGGCGCTCGGCAAAGACCCGGCGGTCGGATTGC 208

Qy        197 AGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAG 256
             |||    ||  | |||||||| ||||| || ||  | || ||||||||  | ||
Db        209 CCGTCGCCTGCGCGATTGAAATGATCCATACGTACTCTTTGATCCATGATGATTTGCCGA 268

Qy        257 CGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGT 316
              ||||| || |||||||  || || || ||    || || |||||||| |  || |||
Db        269 GCATGGACAACGATGATTTGCGGCGCGGCAAGCCGACGAACCATAAAGTGTTCGGCGAGG 328

Qy        317 GGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTC----- 371
              ||  || || || || || || |  || || || |  || |||  |  | |  |
Db        329 CGATGGCCATCTTGGCGGGGGACGGGTTGTTGACGTACGCGTTTCAATTGATCACCGAAA 388

Qy        372 -AAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAG 430
              |||||  |  |  ||  |  ||      |   |  |||||||| | | ||
Db        389 TCGACGATGAGCGCATCCCTCCTTCCGTCCGGCTTCGGCTCATCGAACGGCTGGCGAAAG 448

Qy        431 CAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAAC 490
              |   ||||  |  || ||||||| ||||||      ||||||  || | || ||  ||
Db        449 CGGCCGGTCCGGAAGGGATGGTCGCCGGTCAGGCAGCCGATATGGAAGGAGAGGGGAAAA 508

Qy        491 CAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTT 550
              |  |   |||       | |||   || ||       ||||| ||   | |  |    |
Db        509 CGCTGACGCTTTCGGAGCTCGAATACATTCATCGGCATAAAACCGGGAAAATGCTGCAAT 568

Qy        551 TTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAG 610
              ||       |  ||    |||      |  ||||         |       | | |
Db        569 ACAGCGTGCACGCCGGCGCCTTGATCGGCGGCGCTGATGCCCGGCAAACGCGGGAGCTTG 628

Qy        611 AAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCT 670
              |  |     ||| ||||  |    || || |||   |||||| |  | ||
Db        629 ACGAATTCGCCGCCCATCTAGGCCTTGCCTTTCAAATTCGCGATGATATTCTCGATATTG 688

Qy        671 ATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTA 730
              | || |  |||| |||  | || ||    || |||||||| |     || || |||   |
Db        689 AAGGGGCAGAAGAAAAAATCGGCAAGCCGGTCGGCAGCGACCAAAGCAACAACAAAGCGA 748

Qy        731 CGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGACG 790
              ||||       ||  |   |   | |||||  | || || ||| | |  ||||  || |
Db        749 CGTATCCAGCGTTGCTGTCGCTTGCCGGCGCGAAGGAAAAGTTGGCGTTCCATATCGAGG 808

Qy        791 CAGC 794
              | ||
Db        809 CGGC 812


RESULT 11
```

```
US-09-367-528A-4
; Sequence 4, Application US/09367528A
; Patent No. 6395525
; GENERAL INFORMATION:
;  APPLICANT:  TOYOTA JIDOSHA KABUSHIKI KAISHA
;  TITLE OF INVENTION:  Geranyl Diphosphate Synthetase Gene
;  FILE REFERENCE:  PH-586
;  CURRENT APPLICATION NUMBER: US/09/367,528A
;  CURRENT FILING DATE:  1999-08-16
;  PRIOR APPLICATION NUMBER:  JP97/346686
;  PRIOR FILING DATE:  1997-12-16
;  NUMBER OF SEQ ID NOS:  6
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 894
;   TYPE: DNA
;   ORGANISM: Bacillus stearothermophilus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(894)
US-09-367-528A-4

  Query Match             18.1%;  Score 155.6;  DB 4;  Length 894;
  Best Local Similarity   50.9%;  Pred. No. 3.7e-29;
  Matches  399;  Conservative    0;  Mismatches  379;  Indels    6;  Gaps    1;

Qy         17 TGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTAA 76
              | ||  |   || | |  ||  | |  |   | ||   |   ||| |    | || |
Db         29 TCAACGAGCAAAAACAGGCGGTGGAAACAGCGCTCTCCCGTTATATAGAGCGCTTAGAAG 88

Qy         77 TGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCC 136
              |    |  ||||  || |    |||  ||| ||||  | || || || ||||| ||||
Db         89 GGCCGGCGAAGCTGAAAAAGGCGATGGCGTACTCATTGGAGGCCGGCGGCAAACGAATCC 148

Qy        137 GACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGA 196
              | ||  | ||| |  |   |||  |     | ||    |  ||    | | | ||  ||
Db        149 GTCCGTTGCTGCTTCTGTCCACCGTTCGGGCGCTCGGCAAAGACCCGGCGGTCGGATTGC 208

Qy        197 AGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAG 256
             |||    ||  | |||||||| ||||| || ||  | || ||||||||  | ||
Db        209 CCGTCGCCTGCGCGATTGAAATGATCCATACGTACTCTTTGATCCATGATGATTTGCCGA 268

Qy        257 CGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGT 316
              ||||| || |||||||  || || || ||    || || |||||||| |  || |||
Db        269 GCATGGACAACGATGATTTGCGGCGCGGCAAGCCGACGAACCATAAAGTGTTCGGCGAGG 328

Qy        317 GGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTC----- 371
              ||  || || || || || || |  || || || |  || |||  |  | |  |
Db        329 CGATGGCCATCTTGGCGGGGGACGGGTTGTTGACGTACGCGTTTCAATTGATCACCGAAA 388

Qy        372 -AAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAG 430
              |||||  |  |  ||  |  ||      |   |  |||||||| | | ||
Db        389 TCGACGATGAGCGCATCCCTCCTTCCGTCCGGCTTCGGCTCATCGAACGGCTGGCGAAAG 448

Qy        431 CAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAAC 490
              |   ||||  |  || ||||||| ||||||      ||||||  || | || ||  ||
Db        449 CGGCCGGTCCGGAAGGGATGGTCGCCGGTCAGGCAGCCGATATGGAAGGAGAGGGGAAAA 508

Qy        491 CAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTT 550
              |  |   |||       | |||   || ||       ||||| ||   | |  |    |
Db        509 CGCTGACGCTTTCGGAGCTCGAATACATTCATCGGCATAAAACCGGGAAAATGCTGCAAT 568

Qy        551 TTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAG 610
              ||       |  ||    |||      |  ||||         |       | | |
Db        569 ACAGCGTGCACGCCGGCGCCTTGATCGGCGGCGCTGATGCCCGGCAAACGCGGGAGCTTG 628

Qy        611 AAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCT 670
              |  |     ||| ||||  |    || || |||   |||||| |  | ||
Db        629 ACGAATTCGCCGCCCATCTAGGCCTTGCCTTTCAAATTCGCGATGATATTCTCGATATTG 688

Qy        671 ATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTA 730
              | || |  |||| |||  | || ||    || |||||||| |     || || |||   |
Db        689 AAGGGGCAGAAGAAAAAATCGGCAAGCCGGTCGGCAGCGACCAAAGCAACAACAAAGCGA 748
```

```
Qy        731 CGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGACG 790
              ||||       ||  |   |   |  |||||  | || || ||| | |  ||||  || |
Db        749 CGTATCCAGCGTTGCTGTCGCTTGCCGGCGCGAAGGAAAAGTTGGCGTTCCATATCGAGG 808

Qy        791 CAGC 794
              | ||
Db        809 CGGC 812


RESULT 12
US-08-534-910B-1
; Sequence 1, Application US/08534910B
; Patent No. 5766911
;  GENERAL INFORMATION:
;    APPLICANT:  KOIKE, Ayumi
;    APPLICANT:  OBATA, Shusei
;    APPLICANT:  NISHINO, Tokuzo
;    APPLICANT:  OHNUMA, Shinichi
;    APPLICANT:  NAKAZAWA, Takeshi
     APPLICANT:  OGURA, Kyozo
     APPLICANT:  KOYAMA, Tanetoshi
;    TITLE OF INVENTION:  Mutated Farnesyldiphosphate Synthase Capable
;    TITLE OF INVENTION:  Of Synthesizing Geranylgeranyldiphosphate And Gene Coding Ther
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenyon & Kenyon
;      STREET:  1025 Connecticut Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  U.S.
;      ZIP:  20036-5405
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.25" Floppy Disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;      SOFTWARE:  IBM/Word Perfect 6.1 Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/534,910B
;      FILING DATE:  28-SEPT-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-25253
;      FILING DATE:  14-FEB-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Toffenetti, Judith L.
;      REGISTRATION NUMBER:  39,048
;      REFERENCE/DOCKET NUMBER:  77670/398
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)429-1776
;      TELEFAX:  (202)429-0796
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  894 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Bacillus stearothermophilus
US-08-534-910B-1

  Query Match             17.9%;  Score 154;  DB 1;  Length 894;
  Best Local Similarity   50.8%;  Pred. No. 8.9e-29;
  Matches  398;  Conservative    0;  Mismatches  380;  Indels    6;  Gaps    1;

Qy         17 TGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTAA 76
              | ||  |   || | | |  ||  | | |   | ||   |   |||  |    | || |
Db         29 TCAACGAGCAAAAACAGGCGGTGGAAACAGCGCTCTCCCGTTATATAGAGCGCTTAGAAG 88

Qy         77 TGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCC 136
                 |    | ||||  || |    |||  ||| |||||  | || || || ||||| ||||
Db         89 GGCCGGCGAAGCTGAAAAAGGCGATGGCGTACTCATTGGAGGCCGGCGGCAAACGAATCC 148

Qy        137 GACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGA 196
              | ||  | |||  |   |   |||  |    | ||    |  ||    | | | ||  ||
Db        149 GTCCGTTGCTGCTTCTGTCCACCGTTCGGGCGCTCGGAAAAGACCCGGCGGTCGGATTGC 208

Qy        197 AGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAG 256
                  |||    ||  | ||||||||| |||||  | ||  | || ||||||||  | ||
Db        209 CCGTCGCCTGCGCGATTGAAATGATCCATACGCACTCTTTGATCCATGATGATTTGCCGA 268

Qy        257 CGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGT 316
              |||||  || |||||||  || || || ||    || || |||||||| |  || |||
Db        269 GCATGGACAACGATGATTTGCGGCGCGGCAAGCCGACGAACCATAAAGTGTTCGGCGAGG 328

Qy        317 GGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTC----- 371
              ||  || || || || || || || |  || || ||  | || ||| || | ||  |
Db        329 CGATGGCCATCTTGGCGGGGGACGGGTTGTTGACGTACGCGTTTCAATTGATCACCGAAA 388

Qy        372 -AAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAG 430
               |||||  |  |  ||  |  ||      |      |  |  |||||||| | | ||
Db        389 TCGACGATGAGCGCATCCCTCCTTCCGTCCGGCTTCGGCTCATCGAACGGCTGGCGAAAG 448

Qy        431 CAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAAC 490
              |    ||||  |  || ||||||| ||||||       ||||||  || | || ||  ||
Db        449 CGGCCGGTCCGGAAGGGATGGTCGCCGGTCAGGCAGCCGATATGGAAGGAGAGGGGAAAA 508

Qy        491 CAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTT 550
              |  |    |||       | |||   || ||       ||||| ||   | |  |    |
Db        509 CGCTGACGCTTTCGGAGCTCGAATACATTCATCGGCATAAAACCGGGAAAATGCTGCAAT 568

Qy        551 TTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAG 610
                 ||       |  ||    ||||      |  ||||        |      |  | |
Db        569 ACAGCGTGCACGCCGGCGCCTTGATCGGCGGCGCTGATGCCCGGCAAACGCGGGAGCTTG 628

Qy        611 AAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCT 670
              |   |      ||| |||| |    || || |||   |||||| |  | ||
Db        629 ACGAATTCGCCGCCCATCTAGGCCTTGCCTTTCAAATTCGCGATGATATTCTCGATATTG 688

Qy        671 ATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTA 730
              | || |  |||| |||  | || ||    || |||||||| |     || || |||   |
Db        689 AAGGGGCAGAAGAAAAAATCGGCAAGCCGGTCGGCAGCGACCAAAGCAACAACAAAGCGA 748

Qy        731 CGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGACG 790
              ||||       ||  |   |   |  |||||  | || || ||| | |  ||||  || |
Db        749 CGTATCCAGCGTTGCTGTCGCTTGCCGGCGCGAAGGAAAAGTTGGCGTTCCATATCGAGG 808

Qy        791 CAGC 794
              | ||
Db        809 CGGC 812


RESULT 13
US-09-367-528A-2
; Sequence 2, Application US/09367528A
; Patent No. 6395525
; GENERAL INFORMATION:
;  APPLICANT:  TOYOTA JIDOSHA KABUSHIKI KAISHA
;  TITLE OF INVENTION:  Geranyl Diphosphate Synthetase Gene
;  FILE REFERENCE:  PH-586
;  CURRENT APPLICATION NUMBER: US/09/367,528A
;  CURRENT FILING DATE:  1999-08-16
;  PRIOR APPLICATION NUMBER:  JP97/346686
;  PRIOR FILING DATE:  1997-12-16
;  NUMBER OF SEQ ID NOS:  6
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 894
;   TYPE: DNA
;   ORGANISM: Bacillus stearothermophilus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(894)
US-09-367-528A-2
```

```
  Query Match             17.9%;  Score 154;  DB 4;  Length 894;
  Best Local Similarity   50.8%;  Pred. No. 8.9e-29;
  Matches  398;  Conservative    0;  Mismatches  380;  Indels    6;  Gaps    1;

Qy         17 TGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTAA 76
              | ||  |   || | | |  ||  | |  |   | ||   |   ||| |    | || |
Db         29 TCAACGAGCAAAAACAGGCGGTGGAAACAGCGCTCTCCCGTTATATAGAGCGCTTAGAAG 88

Qy         77 TGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCC 136
              |    |  ||||  || |    |||  ||| |||||  | || || || ||||| ||||
Db         89 GGCCGGCGAAGCTGAAAAAGGCGATGGCGTACTCATTGGAGGCCGGCGGCAAACGAATCC 148

Qy        137 GACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGA 196
              | ||  | ||| |  |   |||  |     | ||    |  ||    | | | ||  ||
Db        149 GTCCGTTGCTGCTTCTGTCCACCGTTCGGGCGCTCGGCAAAGACCCGGCGGTCGGATTGC 208

Qy        197 AGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAG 256
                 |||    ||  | |||||||| ||||| || |   | || ||||||||  | ||
Db        209 CCGTCGCCTGCGCGATTGAAATGATCCATACGTACTTCTTGATCCATGATGATTTGCCGA 268

          257 CGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGT 316
              ||||| || |||||||  || || || ||    || || |||||||| |  || |||
Db        269 GCATGGACAACGATGATTTGCGGCGCGGCAAGCCGACGAACCATAAAGTGTTCGGCGAGG 328

Qy        317 GGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTC----- 371
              ||  || || || || || || | |  || || ||  | || ||| || | ||  |
Db        329 CGATGGCCATCTTGGCGGGGGACGGGTTGTTGACGTACGCGTTTCAATTGATCACCGAAA 388

Qy        372 -AAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAG 430
               |||||  |  |  ||  |  ||      |    |  |  |||||||| | | ||
Db        389 TCGACGATGAGCGCATCCCTCCTTCCGTCCGGCTTCGGCTCATCGAACGGCTGGCGAAAG 448

Qy        431 CAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAAC 490
              |    ||||  |  || ||||||| ||||||        ||||| || | || ||  ||
Db        449 CGGCCGGTCCGGAAGGGATGGTCGCCGGTCAGGCAGCCGATATGGAAGGAGAGGGGAAAA 508

Qy        491 CAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTT 550
              |  |   |||      | |||   || ||      ||||| ||   | |  |    |
Db        509 CGCTGACGCTTTCGGAGCTCGAATACATTCATCGGCATAAAACCGGGAAAATGCTGCAAT 568

Qy        551 TTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAG 610
                 ||      |  ||    ||||     |  ||||         |      |  | |
Db        569 ACAGCGTGCACGCCGGCGCCTTGATCGGCGGCGCTGATGCCCGGCAAACGCGGGAGCTTG 628

Qy        611 AAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCT 670
              |    |         ||| ||||   |    || || |||   ||||||  |  | ||
Db        629 ACGAATTCGCCGCCCATCTAGGCCTTGCCTTTCAAATTCGCGATGATATTCTCGATATTG 688

Qy        671 ATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTA 730
              | || |  |||| |||  | || ||    || |||||||| |     || || |||   |
Db        689 AAGGGGCAGAAGAAAAAATCGGCAAGCCGGTCGGCAGCGACCAAAGCAACAACAAAGCGA 748

Qy        731 CGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGACG 790
              ||||       ||  |   |   | |||||  | || || ||| | |  |||| || |
Db        749 CGTATCCAGCGTTGCTGTCGCTTGCCGGCGCGAAGGAAAAGTTGGCGTTCCATATCGAGG 808

Qy        791 CAGC 794
              | ||
Db        809 CGGC 812


RESULT 14
US-09-217-609A-28
; Sequence 28, Application US/09217609A
; Patent No. 6071733
;  GENERAL INFORMATION:
;    APPLICANT:  MURAMATSU, Masayoshi
;    APPLICANT:  KOIKE, Ayumi
;    APPLICANT:  OGURA, Kyozo
;    APPLICANT:  KOYAMA, Tanetoshi
;    APPLICANT:  SHIMIZU, Naoto
;    APPLICANT:  CHO, Yenwin
;    TITLE OF INVENTION:  Prenyl Diphosphate Synthetase Genes
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenyon & Kenyon
;      STREET:  1025 Connecticut Avenue, NW - Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  US
;      ZIP:  20036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;      SOFTWARE:  WordPerfect 6.1 for Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/217,609A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/873,235
;      FILING DATE:  11-Jun-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  TOFFENETTI, Judith L.
;      REGISTRATION NUMBER:  39,048
;      REFERENCE/DOCKET NUMBER:  10235/1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-429-1776
;      TELEFAX:  202-429-0796
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  486 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  genomic DNA
US-09-217-609A-28

  Query Match             17.5%;  Score 150.6;  DB 3;  Length 486;
  Best Local Similarity   58.4%;  Pred. No. 5.2e-28;
  Matches  277;  Conservative    0;  Mismatches  194;  Indels    3;  Gaps    1;

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              || || || |||||  ||||| | ||  |  |    |||||||||    ||  |  | |
Db          1 GGTGGCAAGCGCATTAGACCATTACTTGTTCTGACTACTTTAGATAGTTTAGGTGGCAAT 60

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
                    |  ||| |       | ||||| || ||||||||||||||| |||||  | |||
Db         61 GCACATGACGGTTTACCATTTGGCATTGCGCTTGAAATGATTCATACGTATTCTTTAATT 120

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              || |||||| | || || |||||||||||||| ||||| || || ||| | || ||||||
Db        121 CACGATGACTTGCCGGCAATGGATAATGATGACTATCGTCGCGGTAAACTCACGAATCAT 180

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||    | || |||    || || ||| | || || |||||  ||  | ||| | || |||
Db        181 AAGCGTTTTGATGAAGCAACAGCTATACTCGCTGGAGATGCATTGCTCACTGATGCTTTT 240

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||   |||| ||   ||       |||| |  ||| | ||| ||  | |  | |
Db        241 CAATGCATTTTAA---ATACGCAGTTAAACGCAGAAATTAAATTATCATTGATTAATTTA 297

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              |  |  ||   |||   | ||| |||||   ||| |||||| | |||||||| |
Db        298 TTAAGTACTGCTTCTGGATCTAATGGCATGGTTTACGGCCAAATGCTCGATATGCAAGGT 357

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              |||    || || |      |  |      |||||   || |||| |   ||||| || |
Db        358 GAACATAAAACATTGACATTAAATGAACTGGAACGTATTCACATACATAAAACCGGTGAN 417

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACA 594
              ||  |   |  ||  || | |||||||     ||    | | |   ||||  ||
Db        418 TTGATTCGTGCAGCANTTGTAAGTGCAGGTATCATANTGANTTTTANTGATGCA 471
```

```
RESULT 15
US-08-873-235B-28
; Sequence 28, Application US/08873235B
; Patent No. 6174715
;  GENERAL INFORMATION:
;    APPLICANT:  MURAMATSU, Masayoshi
;    APPLICANT:  KOIKE, Ayumi
;    APPLICANT:  OGURA, Kyozo
;    APPLICANT:  KOYAMA, Tanetoshi
;    APPLICANT:  SHIMIZU, Naoto
;    APPLICANT:  CHO, Yenwin
;    TITLE OF INVENTION:  Prenyl Diphosphate Synthetase Genes
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenyon & Kenyon
;      STREET:  1025 Connecticut Avenue, NW - Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  US
;      ZIP:  20036
     COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;      SOFTWARE:  WordPerfect 6.1 for Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/873,235B
;      FILING DATE:  11-Jun-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 154441/1996
;      FILING DATE:  14-Jun-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  TOFFENETTI, Judith L.
;      REGISTRATION NUMBER:  39,048
;      REFERENCE/DOCKET NUMBER:  10235/1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-429-1776
;      TELEFAX:  202-429-0796
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  486 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  genomic DNA
US-08-873-235B-28

  Query Match             17.5%;  Score 150.6;  DB 4;  Length 486;
  Best Local Similarity   58.4%;  Pred. No. 5.2e-28;
  Matches  277;  Conservative    0;  Mismatches  194;  Indels    3;  Gaps    1;

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              || || || |||||  |||||  | ||  |  |    |||||||||    ||  |  | |
Db          1 GGTGGCAAGCGCATTAGACCATTACTTGTTCTGACTACTTTAGATAGTTTAGGTGGCAAT 60

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
                  |   |||  |       | |||||  || ||||||||||||||| |||||  | |||
Db         61 GCACATGACGGTTTACCATTTGGCATTGCGCTTGAAATGATTCATACGTATTCTTTAATT 120

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              || |||||| | || || |||||||||||||| ||||| || || ||| | || ||||||
Db        121 CACGATGACTTGCCGGCAATGGATAATGATGACTATCGTCGCGGTAAACTCACGAATCAT 180

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||    | || |||    || || ||| | || || ||||| ||  | ||| | || |||
Db        181 AAGCGTTTTGATGAAGCAACAGCTATACTCGCTGGAGATGCATTGCTCACTGATGCTTTT 240

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
               ||   |||| ||   ||      ||||  |  ||| | ||| ||  | |  |  |
Db        241 CAATGCATTTTAA---ATACGCAGTTAAACGCAGAAATTAAATTATCATTGATTAATTTA 297

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
```

```
                |    |  ||   |||   |  ||| |||||   ||| |||||| | |||||||| |
Db        298 TTAAGTACTGCTTCTGGATCTAATGGCATGGTTTACGGCCAAATGCTCGATATGCAAGGT 357

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              |||    || || |     |  |     |||||   || |||| |   |||| || |
Db        358 GAACATAAAACATTGACATTAAATGAACTGGAACGTATTCACATACATAAAACCGGTGAN 417

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACA 594
              ||  |   |   ||  || | |||||||     ||    | | |   |||| ||
Db        418 TTGATTCGTGCAGCANTTGTAAGTGCAGGTATCATANTGANTTTTANTGATGCA 471


Search completed: May 30, 2003, 17:36:21
Job time : 63 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 30, 2003, 15:24:59 ; Search time 1600 Seconds
(without alignments)
8715.196 Million cell updates/sec

Title: US-09-925-637-63
Perfect score: 861
Sequence: 1 atgacgaatctaccgatgaa.....tattagaaatcgttgattta 861

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

tal number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	151.6	17.6	4484	17	BH770981	BH770981 LLMGtag70
2	89.2	10.4	354	12	BF610197	BF610197 NXSI_055_
3	89.2	10.4	558	9	AL749888	AL749888 AL749888
4	88.6	10.3	657	9	AJ502290	AJ502290 AJ502290
5	88	10.2	425	10	BE323612	BE323612 NF006A09P
6	87.4	10.2	592	12	BG317831	BG317831 NXPV_006_
7	84.6	9.8	355	9	AI780894	AI780894 EST261773
8	84.2	9.8	442	9	AJ302129	AJ302129 AJ302129
9	80	9.3	571	10	BE432698	BE432698 EST399227
c 10	79.2	9.2	475	17	BH439236	BH439236 BOHAC18TF
11	78.6	9.1	468	12	BF519441	BF519441 EST456904
12	78.6	9.1	682	14	BQ856404	BQ856404 QGB3g04.y
13	78.4	9.1	685	13	BM449136	BM449136 DSA032E03
14	77	8.9	636	9	AU238772	AU238772 AU238772
15	75.4	8.8	716	14	BQ506776	BQ506776 EST614191
16	75.2	8.7	863	17	BH584613	BH584613 BOGQH60TR
c 17	75	8.7	770	17	BH250143	BH250143 BOGAG35TF
c 18	74.8	8.7	808	14	BQ506777	BQ506777 EST614192
19	73.4	8.5	384	10	BE521431	BE521431 M19E5STM
20	73	8.5	703	17	BH429531	BH429531 BOHLB93TR
c 21	70.6	8.2	1101	17	CNS0039G	AL063921 Drosophil
22	70.4	8.2	561	12	BF054034	BF054034 EST439264
23	70.4	8.2	630	12	BF270259	BF270259 GA__Eb000
24	70.4	8.2	670	14	BQ401843	BQ401843 GA__Ed004
25	70.4	8.2	689	12	BG439911	BG439911 GA__Ea000
26	70.4	8.2	892	12	BF276100	BF276100 GA__Eb002
27	70.4	8.2	907	12	BG444779	BG444779 GA__Ea002
28	69.8	8.1	701	10	BE238154	BE238154 894040E07
29	69.6	8.1	666	13	BI717621	BI717621 1031021A0
30	69.6	8.1	675	13	BI723806	BI723806 1031067H0
31	68.8	8.0	668	13	BI924206	BI924206 EST544095
32	68.8	8.0	669	12	BG889977	BG889977 EST515828
33	68.8	8.0	681	10	AW096600	AW096600 EST289780
34	68.8	8.0	712	12	BF521402	BF521402 EST458878
35	68	7.9	668	13	BI726915	BI726915 1031088E0
c 36	67.6	7.9	838	17	BH657017	BH657017 BOMBM36TF
37	67.2	7.8	709	17	BH502634	BH502634 BOHJK46TR
38	67	7.8	669	17	BH731669	BH731669 BOHYD34TF
39	67	7.8	823	10	BE640755	BE640755 Cri2_1_F1
40	66.2	7.7	542	14	BQ696145	BQ696145 NXPV_036_
41	64.2	7.5	662	10	AW725795	AW725795 GA__Ea001
42	64	7.4	828	17	BH501300	BH501300 BOHAP61TF
43	64	7.4	852	17	BH432244	BH432244 BOGPB71TF
44	62	7.2	618	9	AU238793	AU238793 AU238793
45	62	7.2	706	13	BI434201	BI434201 EST536962

ALIGNMENTS

RESULT 1
BH770981
LOCUS BH770981 4484 bp DNA linear GSS 01-MAY-2002
DEFINITION LLMGtag706 MG1363 Random Sequence Tag Library Lactococcus lactis subsp. cremoris genomic, DNA sequence.
ACCESSION BH770981
VERSION BH770981.1 GI:20373938
KEYWORDS GSS.
SOURCE Lactococcus lactis subsp. cremoris.
ORGANISM Lactococcus lactis subsp. cremoris
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
REFERENCE 1 (bases 1 to 4484)
AUTHORS Bolotin,A., Ehrlich,S.D. and Sorokin,A.
TITLE Studies of genomes of dairy bacteria Lactococcus lactis
JOURNAL Sci. Aliments, (2002) In press
COMMENT Contact: Sorokin A
Genetique Microbienne
INRA
CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
Tel: 33 1 34 65 25 16
Fax: 33 1 34 65 25 21
Email: sorokine@jouy.inra.fr
best homologue in strain IL1403 is xseA (94%)
Class: shotgun
High quality sequence start: 30
High quality sequence stop: 4456.
FEATURES Location/Qualifiers
source 1. .4484

```
                     /organism="Lactococcus lactis subsp. cremoris"
                     /strain="MG1363"
                     /db_xref="taxon:1359"
                     /clone_lib="MG1363 Random Sequence Tag Library"
                     /note="Vector: pSGMU2; Site_1: SmaI; Library of
                     chromosomal fragments of L.lactis strain MG1363 was
                     prepared by partial AluI digestion or by sonication."
BASE COUNT      1540 a    637 c   1023 g   1284 t
ORIGIN

  Query Match           17.6%;  Score 151.6;  DB 17;  Length 4484;
  Best Local Similarity  52.6%;  Pred. No. 1.9e-26;
  Matches  378;  Conservative    0;  Mismatches  334;  Indels    6;  Gaps    2;

Qy        106 TATTCATTAAATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGAT 165
              ||||| ||||  || || |||||||| || || ||  | || || ||      ||| ||
Db       3738 TATTCCTTAATGGCGGGCGGTAAACGAATTCGTCCTCTCCTTTTTTTGAATCTTTTGGAA 3797

Qy        166 TCACTAAATACCGAGTAT---GAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATT 222
              |  |  ||  ||  |    |   |  |    ||   || || |||||||||
Db       3798 GCTTTTGATTTGGAAGTTACAACGGCCCATTATCAAGTGGCCGCAGCGCTTGAAATGATT 3857

Qy        223 CATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGA 282
              |||||    ||  | ||||||||||||| | || |||||||| ||||||||||||||| ||
Db       3858 CATACGGGCTCGTTAATTCATGATGACTTGCCTGCGATGGACAATGATGATTATCGTCGT 3917

Qy        283 GGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCT 342
              ||||| ||||| || ||||||  || || |||   ||| || ||  | || || ||| ||
Db       3918 GGAAAGTTAACGAACCATAAAAAATTTGATGAAGCGACAGCAATTCTGGCTGGAGATACT 3977

Qy        343 TTATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAA 402
              |||||     |  | ||| |  || |||||||   |||   ||  | |   ||  ||
Db       3978 TTATTTTTCGACCCTTTTTATGTTCTTTCAAGTAGCGAT---TTGTCGGCGCAAACAATT 4034

Qy        403 ATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAA 462
              ||  | |    |      || |    ||    ||      ||||||||| |  ||||||
Db       4035 GTATCATTGACCCGCGAATTGGCTTATGCCTCGGGCTCGTATGGAATGGTTGCTGGTCAA 4094

Qy        463 ATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACAC 522
              || || ||||||    | |||||  ||    |  | ||      | ||    || ||
Db       4095 ATTTTGGATATGGCTGGTGAAGGTAAAGATTTAACTTTGACTCAGATTGAGCAAATTCAT 4154

Qy        523 AAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAAT 582
              |  |||||| ||   | | |||||||||  |  || ||  ||| | |    | ||  |
Db       4155 CGATTAAAAACGGGTCGACTTTTAACTTTTCCTTTTGTGGCTGCTGGAATCGTAGCCCAA 4214

Qy        583 GTCGATGATACAACTAAAGAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTC 642
                 ||  |   |  ||| |  |           |  |  ||||||  |    ||
Db       4215 AAAACAAATCAAGAAATCGAAAAACTTCGTGTAGTCGGACAAATTTTAGGATTAGCTTTT 4274

Qy        643 CAGATTAAAGATGATTTATTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTG 702
              || |||   |||||  | |||||        |    ||  |  | |||||  |
Db       4275 CAAATTCGTGATGACATTTTAGATGTGACGGCCACTTTTGCTGAACTGGGTAAGACGCCT 4334

Qy        703 GGCAGCGATCTTGAAAATAATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCA 762
              || |  ||| ||  | |  | ||    ||.|| ||       |||||    || || ||
Db       4335 GGAAAGGATGTTTTAGAAGAAAAGTCGACTTATGTTGCGCATTTAGGACTTGAAGGAGCT 4394

Qy        763 GAAGATAAATTGACTTATCATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATG 820
              ||     | ||||  || |       | || |    |   ||| ||  |  |  |||
Db       4395 AAAATATCACTGACAGATAAGTTGTCAGAAGTAAAAAAATTACTGACAGAGTTGAATG 4452


RESULT 2
BF610197
LOCUS       BF610197                354 bp    mRNA    linear   EST 14-DEC-2000
DEFINITION  NXSI_055_H03_F NXSI (Nsf Xylem Side wood Inclined) Pinus taeda cDNA
            clone NXSI_055_H03 5', mRNA sequence.
ACCESSION   BF610197
VERSION     BF610197.1  GI:11778609
KEYWORDS    EST.
SOURCE      loblolly pine.
  ORGANISM  Pinus taeda
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE   1  (bases 1 to 354)
  AUTHORS   Sederoff,R.
  TITLE     Molecular Basis of Wood Formation in the Pine Megagenome
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Johnson, Arthur
            North Carolina State University
            Tel: 919 515 7800
            Fax: 919 515 7801
            Email: ajohnson@unity.ncsu.edu
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .354
                     /organism="Pinus taeda"
                     /strain="Coastal plain loblolly pine from North Carolina"
                     /db_xref="taxon:3352"
                     /clone="NXSI_055_H03"
                     /clone_lib="NXSI (Nsf Xylem Side wood Inclined)"
                     /tissue_type="Xylem"
                     /cell_type="Side"
                     /dev_stage="Juvenile"
                     /lab_host="XL1-Blue"
                     /note="Vector: BlueScript SK; Site_1: Eco RI; Site_2: XhoI
                     ; The library is from early (spring) wood, taken from
                     three six-year old trees (three different genotypes), in
                     the juvenile phase. These trees were induced to form side
                     wood by bending to a 45 degree angle and tying them to the
                     ground. Differentiating xylem was harvested from the sides
                     of the inclined stems, and a mixture of all three
                     genotypes was used for the library. oligo-dT primed cDNA
                     was directionally cloned into the EcoRI-XhoI BlueScript SK
                     vector arms. NOTE: The sequences contain a 'cDNA adapter'
                     between the EcoRI site and the start of the EST. The
                     adapter sequence is 'AATTCGGCACGAG'."
BASE COUNT       72 a     68 c     98 g    104 t     12 others
ORIGIN

  Query Match           10.4%;  Score 89.2;  DB 12;  Length 354;
  Best Local Similarity  63.6%;  Pred. No. 1.9e-11;
  Matches  136;  Conservative    0;  Mismatches   78;  Indels    0;  Gaps    0;

Qy        182 ATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTC 241
              | ||| | |  |||       ||   |||| | || || || || |||   ||  | ||||
Db         49 AGGAGATGGCCATGCCAGCTGCCTGTGCAATGGAGATTATCCACACAATGTCTTTGATTC 108

Qy        242 ATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATA 301
              |||||||| | ||  | ||||||||||||||||  || ||||| ||    ||||| || |
Db        109 ATGATGACTTGCCTTCCATGGATAATGATGATTTACGCCGAGGTAAGCCTACAAACCACA 168

Qy        302 AAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTG 361
              | || | |||||||| |||||  |  | ||||||||||| |  |  |    || ||||
Db        169 AGGTCTTTGGTGAGGGCACTGCAGTTCTTGCAGGTGATGCTCTTCTTTCGTTTGCTTTTG 228

Qy        362 AACTTATTTCAAGTGATGATAGATTAACTGATGA 395
              | |  ||| ||   | |   ||    ||||  ||
Db        229 AGCACATTGCAGTGGCTACAAGCAAGACTGTGGA 262


RESULT 3
AL749888
LOCUS       AL749888                558 bp    mRNA    linear   EST 20-JUN-2002
DEFINITION  AL749888 AS Pinus pinaster cDNA clone AS02C09 similar to
            GERANYLGERANYL DIPHOSPHATE SYNTHASE, mRNA sequence.
ACCESSION   AL749888
VERSION     AL749888.1  GI:21491122
KEYWORDS    EST.
SOURCE      Pinus pinaster.
  ORGANISM  Pinus pinaster
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE   1  (bases 1 to 558)
  AUTHORS   Frigerio,J. and Plomion,C.
```

```
  TITLE     Identification of water-deficit responsive genes in Maritime pine
            (Pinus pinaster Ait.) using an EST approach
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Frigerio JM
            Genetique et Amelioration 69
            INRA
            route d'Arcachon 33612 Cestas CEDEX France
            Email: Frigerio@pierroton.inra.fr
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .558
                     /organism="Pinus pinaster"
                     /db_xref="taxon:71647"
                     /clone="AS02C09"
                     /clone_lib="AS"
                     /tissue_type="shoot"
                     /dev_stage="6 weeks old seedling"
                     /lab_host="SOLR"
                     /note="Vector: Uni-ZAP XR; ecotype: Landes; The library
                     was made from the aerial part (above the collar) of 6
                     weeks old seedlings grown in hydroponic conditions. A
                     three weeks drought stress treatment was applied by
                     lowering the osmotic potential of the nutrient solution to
                     -0.45 MPa using 170 g/l of polyethylene glycol as an
                     osmoticum. A mixture of genotypes were used. Oligo-dT
                     primed cDNA was directionally cloned into the EcoRI-XhoI
                     lambda-ZAP vector arms and mass-excised to form a
                     pBluescript phagemid"
BASE COUNT      158 a    110 c    137 g    153 t
ORIGIN

  Query Match             10.4%;  Score 89.2;  DB 9;  Length 558;
  Best Local Similarity   57.6%;  Pred. No. 2.1e-11;
  Matches  160;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy         92 AAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTAT 151
              | |||   |||  ||||||  |    ||||| || || ||  |  ||||  |  | |
Db        281 ATGAAGCAATGAGGTATTCGCTTCTGGCTGGGGGAAAGCGTGTGAGACCCATATTATGCA 340

Qy        152 TACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCAC 211
              |    |||  ||     ||  |       | || | ||  |||      ||   ||||
Db        341 TCGCTTCTTGCGAGCTCGTAGGTGGGTCTGAGGAATCAGTCATGCCATCGGCCTGTGCAA 400

Qy        212 TAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATG 271
              | || ||||||||  ||    || || ||||||||||  | ||    ||||| ||||| |
Db        401 TTGAGATGATTCACACTATGTCTCTGATTCATGATGATTTGCCTTGTATGGACAATGACG 460

Qy        272 ATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAG 331
              |||   || |||| ||    |||||||| || || | ||| |||     |||  |  | |
Db        461 ATTTGAGAAGAGGGAAGCCCACAAATCACAAGGTGTTTGGGGAGGATGTTGCTGTTCTGG 520

Qy        332 CAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATT 369
              |||| ||||||  | ||  |    || |||||  |  |||
Db        521 CAGGGGATGCTCTGTTGGCATTTGCCTTTGAGCACATT 558


RESULT 4
AJ502290
LOCUS       AJ502290                657 bp    mRNA    linear   EST 09-AUG-2002
DEFINITION  AJ502290 MTAMP Medicago truncatula cDNA clone mtgmadc120019a01,
            mRNA sequence.
ACCESSION   AJ502290
VERSION     AJ502290.1  GI:22083218
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 657)
  AUTHORS   Manthey,K., Bartelsmeier,V., Baier,M.C., Meyer,F., Bartels,D.,
            Bekel,T., Linke,B., Grunwald,U., Franken,P., Kuester,H., Perlick
            ,A.M. and Puehler,A.
  TITLE     Detection of transcript sequences from mycorrhizal roots of the
            model mycorrhiza Medicago truncatula genotype A17 - Glomus mosseae
            using the approach of an EST genome project
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kuester H
            Lehrstuhl fuer Genetik
            Universitaet Bielefeld
            Postfach 100131, D-33501 Bielefeld, Germany.
FEATURES             Location/Qualifiers
     source          1. .657
                     /organism="Medicago truncatula"
                     /db_xref="taxon:3880"
                     /clone="mtgmadc120019a01"
                     /clone_lib="MTAMP"
                     /tissue_type="mycorrhizal roots"
                     /dev_stage="three weeks"
                     /note="Vector: pGEM-T; Site_1: PstI; Site_2: SphI;
                     genotype A17; cDNA was prepared from polyA+ enriched RNA
                     from mycorrhizal roots harvested after three weeks. The
                     cDNA was directionally ligated by MediGenomix into the
                     pGEM-T vector from Promega using GCATGCGGCCGAGGCGGCCGACATG
                     and CTGCAGGCCATTATGGCCGGG adapters. Plasmids containing
                     cDNA inserts were propagated in E. coli DH10B cells."
BASE COUNT      213 a    167 c    117 g    160 t
ORIGIN

  Query Match             10.3%;  Score 88.6;  DB 9;  Length 657;
  Best Local Similarity   57.3%;  Pred. No. 3.1e-11;
  Matches  160;  Conservative    0;  Mismatches  119;  Indels    0;  Gaps    0;

Qy         86 AGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCGACCAGTTC 145
              || |  | |||  ||||   || ||  |    |||||||| || ||  |||| || || |
Db        345 AGATTTACGAAGCTATGCGATACTCTCTCCTCGCTGGAGGAAAGCGTGTCCGTCCGGTAC 404

Qy        146 TGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAA 205
              | |   |   | | | |       ||      | |  | ||   |   ||       |||
Db        405 TCTGTATTGCCGCCTGCGGACTTGTAGGCGGCAAAGAAGAAGCATCAATTCCATCGGCAT 464

Qy        206 TTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATA 265
                ||  | |||||||| || ||    ||  | || |||||||| |||||    |||||||
Db        465 GCGCTGTGGAAATGATCCACACCATGTCGTTAATCCATGATGATCTACCGTGCATGGATA 524

Qy        266 ATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGA 325
              | || |||    ||||||||||   |||||| || ||||| || ||||||       ||
Db        525 ACGACGATCTCAGACGAGGAAAGCCAACAAACCACAAAGTCTACGGTGAGGACGTAGCTG 584

Qy        326 TATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAAC 364
              |  | || || |||||  || |  |    || |||||||
Db        585 TCCTTGCCGGGGATGCACTACTTTCCTTCGCGTTTGAAC 623


RESULT 5
BE323612
LOCUS       BE323612                425 bp    mRNA    linear   EST 21-DEC-2000
DEFINITION  NF006A09PL1F1065 Phosphate starved leaf Medicago truncatula cDNA
            clone NF006A09PL 5', mRNA sequence.
ACCESSION   BE323612
VERSION     BE323612.2  GI:11967172
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 425)
  AUTHORS   Liu,J., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores
            ,H.R., Inman,J.T., Weller,J.W., May,G.D. and Harrison,M.J.
  TITLE     Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula phosphate-starved leaf library
  JOURNAL   Unpublished (2000)
COMMENT     On Jul 14, 2000 this sequence version replaced gi:9197389.
```

```
            Contact: Harrison MJ
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 221 7325
            Fax: 580 221 7380
            Email: mjharrison@noble.org
            Medicago Genome Initiative accession: MGI:S:20177
            Insert Length: 818   Std Error: 0.00
            Plate: 006  row: A  column: 09
            Seq primer: TCACACAGGAAACAGCTATGAC.
FEATURES              Location/Qualifiers
     source           1. .425
                      /organism="Medicago truncatula"
                      /db_xref="taxon:3880"
                      /clone="NF006A09PL"
                      /clone_lib="Phosphate starved leaf"
                      /tissue_type="leaf"
                      /dev_stage="trifoliate"
                      /note="Vector: Lambda Zap; At the trifoliate stage, M.
                      truncatula plants were transplanted to phosphate-free sand
                      and grown for a further 30 days.  During this 30 day
                      period, the plants were fertilized twice weekly with 1/2
                      Hoaglands solution containing only 20uM potassium
                      phosphate.  RNA was prepared from above ground tissues."
BASE COUNT      109 a    116 c     96 g    104 t
ORIGIN

  Query Match             10.2%;  Score 88;  DB 10;  Length 425;
  Best Local Similarity   57.1%;  Pred. No. 4e-11;
  Matches  160;  Conservative    0;  Mismatches  120;  Indels    0;  Gaps    0;

Qy         92 AAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTAT 151
              | |||   |||   ||  ||  |    || || || |||||| | || || ||||| |  |
Db        124 ATGAAGCCATGCGTTACTCTCTTCTCGCCGGCGGAAAACGCGTTCGGCCGGTTCTCTGTT 183

Qy        152 TACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCAC 211
              ||  | | |  ||     |          |   | | |  |||     |||    ||
Db        184 TAGCCGCCTGTGAACTCGTCGGAGGAACCGAACCGATGGCGATGCCAGCCGCCTGCGCCG 243

Qy        212 TAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATG 271
              | ||||||||  ||  ||    ||  ||  ||  ||  ||||||||  ||    ||||||||  ||||
Db        244 TTGAAATGATCCACACGATGTCTCTCATCCACGATGACCTTCCTTGTATGGATAACGATG 303

Qy        272 ATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAG 331
              ||   ||  |||||  |||   |||||  ||  |||||  |   ||  ||       |||   |  | |
Db        304 ATCTCCGGCGAGGTAAACCTACAAACCACAAAGTCTTCGGAGAAGACGTTGCTGTTCTCG 363

Qy        332 CAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTC 371
              | ||  ||||||  |   |   ||    ||  ||  ||||  ||||  |
Db        364 CCGGAGATGCTCTTCTCGCTTTTGCTTTCGAACATATTGC 403


RESULT 6
BG317831
LOCUS       BG317831                592 bp    mRNA    linear   EST 26-FEB-2001
DEFINITION  NXPV_006_B05_F NXPV (Nsf Xylem Planings wood Vertical) Pinus taeda
            cDNA clone NXPV_006_B05 5', mRNA sequence.
ACCESSION   BG317831
VERSION     BG317831.1  GI:13127261
KEYWORDS    EST.
SOURCE      loblolly pine.
  ORGANISM  Pinus taeda
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE   1  (bases 1 to 592)
  AUTHORS   Sederoff,R.
  TITLE     Molecular Basis of Wood Formation in the Pine Megagenome
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Johnson, Arthur
            North Carolina State University
            Tel: 919 515 7800
            Fax: 919 515 7801
            Email: ajohnson@unity.ncsu.edu
            Seq primer: T3.
FEATURES              Location/Qualifiers
     source           1. .592
                      /organism="Pinus taeda"
                      /strain="Coastal plain loblolly pine from North Carolina"
                      /db_xref="taxon:3352"
                      /clone="NXPV_006_B05"
                      /clone_lib="NXPV (Nsf Xylem Planings wood Vertical)"
                      /tissue_type="Xylem"
                      /cell_type="Planings (secondary)"
                      /dev_stage="Transitional"
                      /lab_host="XL1-Blue"
                      /note="Vector: BlueScript SK; Site_1: Eco RI; Site_2: XhoI
                      ; The library is from early (spring) secondary wood, taken
                      from a ten year old tree in the transitional phase. The
                      tree is a kind gift of the Westvaco Corporation. Secondary
                      xylem was harvested from the tree by peeling back the bark
                      and primary xylem and then removing the underlying tissue
                      with a block plane. NOTE: The sequences contain a 'cDNA
                      adapter' between the EcoRI site and the start of the EST.
                      The adapter sequence is 'AATTCGGCACGAG'."
BASE COUNT      163 a    118 c    134 g    154 t     23 others
ORIGIN

  Query Match             10.2%;  Score 87.4;  DB 12;  Length 592;
  Best Local Similarity   53.9%;  Pred. No. 5.9e-11;
  Matches  151;  Conservative    0;  Mismatches  129;  Indels    0;  Gaps    0;

Qy         92 AAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTAT 151
              | |||   |||  ||||||  |    ||||| || || ||  |  ||||  |  | |
Db        304 ATGAAGCAATGAGGTATTCGCTTCTGGCTGGGGGAAAGCGTGTGAGACCCATATTATGCA 363

Qy        152 TACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCAC 211
              |     |||  |||    ||  |        | || | ||  |||      ||   ||||
Db        364 TCGCTGCTTGCGATCTCGTAGGTGGGTCTGAGGAATCAGTCATGCCATCGGCCTGTGCAA 423

Qy        212 TAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATG 271
              | || ||||||||  |||   || || ||||||||||  | ||    |||||  ||||| |
Db        424 TTGAGATGATTCACACAATGTCTCTGATTCATGATGATTTGCCTTGTATGGACAATGACG. 483

Qy        272 ATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAG 331
              |||   || | || ||    ||||||||  |    | ||| |         ||      | |
Db        484 ATTTGAGAAGNGGGAAGCCCACAAATCACNANNNGTTTGGGGNNNATGNNGCTGNNNTNG 543

Qy        332 CAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTC 371
              | |   |||||||  |   ||   ||  ||  | |  ||| |
Db        544 CNGNNNATGCTTTGTNNNCTTTTGCCTTNNAGCACATTGC 583


RESULT 7
AI780894
LOCUS       AI780894                355 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST261773 tomato susceptible, Cornell Lycopersicon esculentum cDNA
            clone cLES13B2, mRNA sequence.
ACCESSION   AI780894
VERSION     AI780894.1  GI:5278935
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 355)
  AUTHORS   D' Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,
            Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman
            ,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Tanksley,S.D.,
            Giovannoni,J.J. and Martin,G.B.
  TITLE     Generation of ESTs from Pseudomonas susceptible tomato
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
```

```
           Clemson University
           100 Jordan Hall, Clemson, SC 29634, USA
           Email: http://www.genome.clemson.edu/orders/index.html
           5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .355
                     /organism="Lycopersicon esculentum"
                     /cultivar="R11-13 (Rio Grande x Money Maker)"
                     /db_xref="taxon:4081"
                     /clone="cLES13B2"
                     /clone_lib="tomato susceptible, Cornell"
                     /tissue_type="leaf"
                     /dev_stage="4-week old"
                     /lab_host="SOLR"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; cLES - Tomato Pseudomonas Susceptible EST Library.
                     Directionally cloned  cDNAs inserted into pBlueScript SK(-
                     ) at 5' end with EcoRI and 3' end  with XhoI site"
BASE COUNT       98 a     82 c     87 g     88 t
ORIGIN

  Query Match             9.8%;  Score 84.6;  DB 9;  Length 355;
  Best Local Similarity   56.7%;  Pred. No. 2.6e-10;
  Matches  156;  Conservative    0;  Mismatches  119;  Indels    0;  Gaps    0;

Qy         92 AAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTAT 151
              | |||   |||  ||| ||| |    |||||||| |||||  |||| || . | || |
Db         67 ATGAAGCCATGAGGTACTCACTTCTAGCTGGAGGAAAACGTGTCCGGCCGATCCTCTGCA 126

Qy        152 TACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCAC 211
              |     |||  ||     ||      ||  |    |||| |||        ||    |||
Db        127 TGGCTTCTTGCGAAGTTGTAGGAGGGGATGAATCCTTAGCTATACCTGCAGCTTGCGCAG 186

Qy        212 TAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATG 271
              | || ||||| |||||    |||||  | ||||||| || ||    ||||| |   |||
Db        187 TTGAGATGATCCATACCATGTCACTCGTCCATGATGATCTTCCCTGCATGGACAGCTATG 246

Qy        272 ATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAG 331
              ||   || || ||  |    || || |||||  | | ||| ||    |||||  |  | |
Db        247 ATCTACGTCGTGGGGAGCCCACGAACCATAAGATTTTTGGAGAAGACACTGCAGTTCTTG 306

Qy        332 CAGGTGATGCTTTATTAACTAAAGCATTTGAACTT 366
              || | |||||  |  || ||  || |||||||
Db        307 CATGGGATGCACTTATATCTTTGGCCTTTGAACAT 341


RESULT 8
AJ302129
LOCUS       AJ302129                 442 bp    mRNA    linear   EST 03-APR-2001
DEFINITION  AJ302129 Solanum tuberosum cv. Bintje leaf Solanum tuberosum cDNA
            clone S3-2B, mRNA sequence.
ACCESSION   AJ302129
VERSION     AJ302129.1  GI:13536952
KEYWORDS    EST.
SOURCE      potato.
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 442)
  AUTHORS   Beyer,K., Binder,A., Boller,T. and Collinge,M.A.
  TITLE     Identification of potato genes induced during colonization by
            Phytophthora infestans
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Beyer K
            Plant Biology
            Friedrich Miescher-Institut
            Maulbeerstrasse 66, 4058 Basel, Switzerland.
FEATURES             Location/Qualifiers
     source          1. .442
                     /organism="Solanum tuberosum"
                     /cultivar="Bintje"
                     /db_xref="taxon:4113"
                     /clone="S3-2B"
                     /clone_lib="Solanum tuberosum cv. Bintje leaf"
                     /tissue_type="leaf"
                     /note="infected with Phytophthora infestans"
BASE COUNT      104 a    102 c    125 g    111 t
ORIGIN

  Query Match             9.8%;  Score 84.2;  DB 9;  Length 442;
  Best Local Similarity   57.4%;  Pred. No. 3.4e-10;
  Matches  152;  Conservative    0;  Mismatches  113;  Indels    0;  Gaps    0;

Qy        109 TCATTAAATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCA 168
              ||| |    |||||||| |||||  |||| ||  | || |    |     ||| ||
Db          1 TCACTTCTAGCTGGAGGAAAACGTGTCCGGCCGATCCTCTGCATGGCTTCTTGTGAAGTT 60

Qy        169 CTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACA 228
              ||     ||  |    |||| |||        ||    ||  | || ||||| |||||
Db         61 GTAGGAGGGGATGAATCCTTAGCTATACCTGCAGCTTGCGCGGTTGAGATGATCCATACC 120

Qy        229 TATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAA 288
              |||||  | |||||||| || ||    ||||| || ||||||   || || || ||
Db        121 ATGTCACTCGTCCATGATGATCTTCCTTGCATGGACAACGATGATCTACGTCGTGGCAAG 180

Qy        289 TTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTA 348
              || || ||||| || | ||| ||    |||||  |  | | ||| |||||  | ||
Db        181 CCCACGAACCATAAGGTTTTTGGGGAAAACACTGCAGTTCTTGTAGGGGATGCACTTTTG 240

Qy        349 ACTAAAGCATTTGAACTTATTTCAA 373
              ||   || |||||| | |  | |
Db        241 TCTTTGGCCTTTGAACATGTGGCTA 265


RESULT 9
BE432698
LOCUS       BE432698                 571 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST399227 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
            clone cLEG9L24, mRNA sequence.
ACCESSION   BE432698
VERSION     BE432698.1  GI:9430541
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 571)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
            Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
            Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
            ,S.D.
  TITLE     Generation of ESTs from tomato fruit tissue, breaker stage
  JOURNAL   Unpublished (2000)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .571
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEG9L24"
                     /clone_lib="tomato breaker fruit, TIGR"
                     /tissue_type="Pericarp"
                     /dev_stage="breaker"
                     /lab_host="SOLR"
                     /note="Vector: pBluescriptSKmCUadapt; Site_1: EcoR1;
                     Site_2: XhoI; Fruit were harvested at the breaker stage
                     (first sign of lycopene accumulation on the blossom end of
                     the fruit).  Fruit were cut in half and the seeds and
```

locules were discarded prior to freezing the pericarp."
BASE COUNT 159 a 107 c 135 g 170 t
ORIGIN

Query Match 9.3%; Score 80; DB 10; Length 571;
Best Local Similarity 53.1%; Pred. No. 3.8e-09;
Matches 170; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

```
Qy         43 AATGAATTATCGGTTGCGATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATG 102
Db        239 AATAAAGCTTTGGATGAGGCTATAATGGTAAAAGACCCACCTAAGATCCATGAAGCAATG 298

Qy        103 TTGTATTCATTAAATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTA 162
Db        299 CGTTATTCGCTTCTCGCCGGCGGGAAGAGAGTCCGGCCGATGCTCTGTCTTGCTGCCTGT 358

Qy        163 GATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATT 222
Db        359 GAACTTGTTGGGGGAAACCAAGGAAATGCCATGGCGGCTGCTTGTGCTGTTGAGATGATA 418

Qy        223 CATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGA 282
Db        419 CATACTATGTCTCTAATTCATGATGATTTGCCTTGTATGGATGACGACGATCTCCGCCGT 478

Qy        283 GGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCT 342
Db        479 GGAAAGCCGACGAATCATAAAGTGTACGGTGAGGATGTGGCGGTCCTCGCCGGAGATGCG 538

Qy        343 TTATTAACTAAAGCATTTGA 362
Db        539 CTACTTGCTTTCGCATTCGA 558
```

RESULT 10
BH439236/c
LOCUS BH439236 475 bp DNA linear GSS 12-DEC-2001
DEFINITION BOHAC18TF BOHA Brassica oleracea genomic clone BOHAC18, DNA sequence.
ACCESSION BH439236
VERSION BH439236.1 GI:17624950
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 475)
AUTHORS Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other_GSSs: BOHAC18TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.
FEATURES Location/Qualifiers
source 1. .475
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOHAC18"
/clone_lib="BOHA"
/note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT 97 a 163 c 122 g 93 t
ORIGIN

Query Match 9.2%; Score 79.2; DB 17; Length 475;
Best Local Similarity 50.5%; Pred. No. 5.7e-09;
Matches 192; Conservative 0; Mismatches 188; Indels 0; Gaps 0;

```
Qy        106 TATTCATTAAATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGAT 165
Db        474 TACTCTCTCCTCGCCGGCGGGAAGCGAGTGAGACCGGTTCTCTGCATCGCCGCGTGCGAG 415

Qy        166 TCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCAT 225
Db        414 CTGGTCGGCGGCAAAGAATCCGTCGCTTTGCCGGCGGCTTGCGCGGTGGAAATGATCCAC 355

Qy        226 ACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGA 285
Db        354 ACGATGTCGCTGATCCACGACGATCTCCCGTGTATGGACAACGACGACCTCCGCCGCGGA 295

Qy        286 AAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTA 345
Db        294 AAACCCACCAACCACAAAGTGTTCGGCGAAGACGTCGCGGTTTTAGCCGGAGACGCGCTT 235

Qy        346 TTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATA 405
Db        234 TTATCGTTCGCGTTTGAGCACTTGGCGTCGACGTCGGGGGTGGCTCCGGCGAGGGTGGTT 175

Qy        406 AAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATG 465
Db        174 AGAGCGATCGGGGAGCTGGCGAGAGCCGTTGGATCAGAAGGGCTTGTGGCGGGTCAGATT 115

Qy        466 TTAGATATGCAAAGCGAAGG 485
Db        114 GTGGATATCAGCAGCGAAGG 95
```

RESULT 11
BF519441
LOCUS BF519441 468 bp mRNA linear EST 08-DEC-2000
DEFINITION EST456904 DSIL Medicago truncatula cDNA clone pDSIL-20N21, mRNA sequence.
ACCESSION BF519441
VERSION BF519441.1 GI:11608124
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
REFERENCE 1 (bases 1 to 468)
AUTHORS Fedorova,M., Pierson,B.L., Samac,D.A., Vance,C.P., Gantt,G.S., Peng,H., Ellis,L., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S., Holt,I.E. and Fraser,C.M.
TITLE ESTs from leaves of Medicago truncatula after inoculation with Colletotrichum trifolii
JOURNAL Unpublished (2000)
COMMENT Contact: Deborah A. Samac
Department of Plant Pathology
University of Minnesota
495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
Tel: 612 625 1243
Fax: 651 649 5058
Email: debbys@puccini.crl.umn.edu
University of Minnesota name: M276873e
TIGR sequence name: MTFCG83TK
More information is available at: http://chrysie.tamu.edu/medicago
Seq primer: SKmod (CTA gAA CTA gtg gAT CC).
FEATURES Location/Qualifiers
source 1. .468
/organism="Medicago truncatula"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pDSIL-20N21"
/clone_lib="DSIL"
/tissue_type="leaves infected with Colletotrichum trifolii"

/dev_stage="cotyledons and primary leaves harvested 5 and 8 days after inoculation with Colletotrichum trifolii"
/lab_host="E. coli strain XLOLR"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; cDNA was prepared from polyA+ enriched RNA from cotyledons and primary leaves harvested 5 and 8 days after inoculation with Colletotrichum trifolii. The cDNA was directionally ligated into the Uni-ZAP XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-ZAP phage using Ex-Assist helper phage and propagated in XLOLR cells. Note: EST may be of fungal origin."

BASE COUNT 129 a 129 c 93 g 117 t
ORIGIN

Query Match 9.1%; Score 78.6; DB 12; Length 468;
Best Local Similarity 56.9%; Pred. No. 7.9e-09;
Matches 144; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

```
Qy        92 AAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTAT 151
Db       210 ATGAAGCCATGCGTTACTCTCTTCTCGCCGGCGGAAAACGCGTTCGGCCGGTTCTCTGTT 269

Qy       152 TACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCAC 211
Db       270 TAGCCGCCTGTGAACTCGTCGGAGGAACCGAACCGATGGCGATGCCAGCCGCCTGCGCCG 329

Qy       212 TAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATG 271
Db       330 TTGAAATGATCCACACGATGTCTCTCATCCACGATGACCTTCCTTGTATGGATAACGATG 389

Qy       272 ATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAG 331
Db       390 ATCTCCGGCGAGGTAAACCTACAAACCACAAAGTCTTCGGAGAAGACGTTGCTGTTCTCG 449

Qy       332 CAGGTGATGCTTT 344
Db       450 CCGGAGATGCTCT 462
```

RESULT 12
BQ856404
LOCUS BQ856404 682 bp mRNA linear EST 14-AUG-2002
DEFINITION QGB3g04.yg.ab1 QG_ABCDI lettuce salinas Lactuca sativa cDNA clone QGB3g04, mRNA sequence.
ACCESSION BQ856404
VERSION BQ856404.1 GI:22241869
KEYWORDS EST.
SOURCE Lactuca sativa.
ORGANISM Lactuca sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae; Lactuca.
REFERENCE 1 (bases 1 to 682)
AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z., Church,S., Jackson,L. and Bradford,K.
TITLE Lettuce and Sunflower ESTs from the Compositae Genome Project http://compgenomics.ucdavis.edu/
JOURNAL Unpublished (2002)
COMMENT Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
belongs to contig QG_CA_Contig1448, see http://cgpdb.ucdavis.edu/ for details.
Plate: QGB3 row: g column: 04.

FEATURES Location/Qualifiers
source 1. .682
/organism="Lactuca sativa"
/cultivar="Salinas"
/db_xref="taxon:4236"
/clone="QGB3g04"
/clone_lib="QG_ABCDI lettuce salinas"
/lab_host="E.coli"
/note="Vector: pBRcDNASfiAB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5′ and 3′ tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_LIB=QG_ABCDI lettuce salinas
TAG_TISSUE=flowers post-fertilized
TAG_SEQ=TGCCATCGGG"

BASE COUNT 160 a 177 c 200 g 145 t
ORIGIN

Query Match 9.1%; Score 78.6; DB 14; Length 682;
Best Local Similarity 48.4%; Pred. No. 8.6e-09;
Matches 281; Conservative 0; Mismatches 294; Indels 6; Gaps 2;

```
Qy        92 AAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTAT 151
Db        72 ATGAAGCTATGCGTTACTCTCTCCTCGCCGGTGGGAAGCGTGTGAGGCCCATTCTCTGCA 131

Qy       152 TACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCAC 211
Db       132 TCGCCGCTTGCGAGCTCGTCGGAGGTGATGAATCCACCGCCATGCCAGCCGCCTGCGCTG 191

Qy       212 TAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATG 271
Db       192 TTGAAATGATACACACCATGTCGTTGATTCACGACGACCTTCCGTGCATGGATAACGACG 251

Qy       272 ATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAG 331
Db       252 ATTTTCGCCGGGGGAAACCAACAAACCACAAGGTTTACGGCGAAGACGTGGCGGTCCTCG 311

Qy       332 CAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTG---ATGATAGATTAA 388
Db       312 CCGGCGACTCTCTCCTCGCCTTCGCTTTCCAACACATCGCAAGCGCCACCGTCGGCGCGT 371

Qy       389 CTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAA 448
Db       372 CTCCTGAACGCGTTTTGGCCGCCGTCGGAGAACTGGCGAAGTCGATTGGGACTGAAGGGT 431

Qy       449 TGGTCGGCGGTCAAATGTTAGATAT---GCAAAGCGAAGGCCAACCAATTGATCTTGAAA 505
Db       432 TGGTCGCGGGTCAAGTGGTGGACATAGCTTCAACAGGTGCAAAGGACGTCGGACTGGACC 491

Qy       506 CTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTATGAGTG 565
Db       492 AGCTTGAGTTCATCCACATCCACAAGACGGCGGCGCTATTGGAAGCAGCCGTCGTATTGG 551

Qy       566 CAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAGAAAGTTATAGTTATC 625
Db       552 GAGCAATATTAGGTGGAGGAAGCGACACCCAAGTGGAAAAGCTGAGAAAGTTCGCGAGAT 611

Qy       626 ATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGAC 666
Db       612 GCATTGGATTATTGTTTCAGGTGGTTGACGACATATTAGAC 652
```

RESULT 13
BM449136
LOCUS BM449136 685 bp mRNA linear EST 01-APR-2002
DEFINITION DSA032E03_60893 An expressed sequence tag database for the halotolerant green alga, Dunaliella salina Dunaliella salina cDNA

```
            clone DSA032E03 5, mRNA sequence.
ACCESSION   BM449136
VERSION     BM449136.1  GI:19854708
KEYWORDS    EST.
SOURCE      Dunaliella salina.
  ORGANISM  Dunaliella salina
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Dunaliellaceae; Dunaliella.
REFERENCE   1  (bases 1 to 685)
  AUTHORS   Cushman,J.C.
  TITLE     An expressed sequence tag database for the halotolerant green alga,
            Dunaliella salina
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Cushman JC
            Department of Biochemistry
            University of Nevada
            MS200, Reno, NV 89557-0014, USA
            Tel: 775-784-1918
            Fax: 775-784-1650
            Email: jcushman@unr.edu
            PCR PRimers
            FORWARD: T3 20mer
            BACKWARD: T7 21mer
            Plate: 032  row: E  column: 03
            Seq primer: T3 20mer
            High quality sequence stop: 685.
FEATURES             Location/Qualifiers
     source          1. .685
                     /organism="Dunaliella salina"
                     /db_xref="taxon:3046"
                     /clone="DSA032E03"
                     /clone_lib="An expressed sequence tag database for the
                     halotolerant green alga, Dunaliella salina"
                     /tissue_type="Cells, which was adapted in 2.5M NaCI via a
                     incremental series from 1.7 to 2.0 to 2.25 to 2.5 M NaCl,
                     were exposed to 3.4 M NaCl for 5 hours"
                     /cell_type="Green"
                     /note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Library construction was performed
                     according to Stratagene's recommended protocol for the
                     Lambda UniZapXR vector and cDNA synthesis kit."
BASE COUNT      153 a    182 c    219 g    131 t
ORIGIN

  Query Match             9.1%;  Score 78.4;  DB 13;  Length 685;
  Best Local Similarity   48.6%;  Pred. No. 9.6e-09;
  Matches  245;  Conservative    0;  Mismatches  256;  Indels    3;  Gaps    1;

Qy        235 CTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACA 294
              || || |||||||| | ||  | ||||| || ||||| |  ||  ||||      ||
Db         10 CTGATCCATGATGACTTGCCCTCCATGGACAACGATGACTTCCGCAGAGGGCGGCCCACC 69

Qy        295 AATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAA 354
              || || || || || || |||   | ||| ||  | || || |||||  |  |  ||
Db         70 AACCACAAGGTGTACGGAGAGGACATTGCCATCCTTGCTGGAGATGCCCTGCTCTCTTTC 129

Qy        355 GCATTTGAACTTATTTCAAGTG---ATGATAGATTAACTGATGAAGTAAAAATAAAAGTT 411
              || ||||| | | || |  | |       | |      |   |||         |    ||
Db        130 GCCTTTGAGCATGTTGCGCGCGCGACCACTGGCACCTCGCCTGAGCGTGTGCTGCGTGTG 189

Qy        412 CTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGAT 471
               | |    ||||   |  ||   ||||   | |||  ||   || || ||  |  | ||
Db        190 ATTCTGGAGCTGGGCAAGGCTGTTGGTGCAGATGGGCTGACAGGAGGACAGGTTGTGGAC 249

Qy        472 ATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAA 531
              ||  | ||||||  | |     | |  || ||     || |    || ||  ||   ||
Db        250 ATTAAGAGCGAAAACGAGGAGGTGGGCCTGGAGGTGCTGCAGTACATCCATGAACACAAG 309

Qy        532 ACAGGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGAT 591
              |||| |||  |  |         | ||  || |||  ||     | |    ||  || ||
Db        310 ACAGCAGCCCTGCTGGAGGCATCCGTGGTGTGTGGCGCTCTCGTGGGCGGTGCGGACGAC 369

Qy        592 ACAACTAAAGAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAA 651
```

```
              ||    ||  |  |    |  ||        |  | ||  |  |||||| |
Db        370 GTGACAGTGGAGAAGCTGCGCAAGTACGCGCGCAACATTGGCCTCGCGTTCCAGGTGGTT 429

Qy        652 GATGATTTATTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGAT 711
              || ||| |  | || |||           |  | | | || || | |  |||| ||
Db        430 GACGATATCCTGGATTGCACCCAGACCACGGAGATGCTGGGCAAGACGGCAGGCAAGGAC 489

Qy        712 CTTGAAAATAATAAAAGTACGTAC 735
               ||||    || || |  ||||||
Db        490 ATTGATGTGAACAAGACCACGTAC 513


RESULT 14
AU238772
LOCUS       AU238772                636 bp    mRNA    linear   EST 01-APR-2002
DEFINITION  AU238772 RAFL19 Arabidopsis thaliana cDNA clone RAFL19-12-M11 5',
            mRNA sequence.
ACCESSION   AU238772
VERSION     AU238772.1  GI:19877941
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 636)
  AUTHORS   Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
            Akiyama,K., Enju,A., Oono,Y., Sakurai,T., Carninci,P., Kawai,J.,
            Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu
            ,M., Hayashizaki,Y. and Shinozaki,K.
  TITLE     Large scale analysis of Arabidopsis full-length cDNA
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Motoaki Seki
            Plant Functional Genomics Research Group
            RIKEN Genomic Sciences Center
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: 81-298-36-4359
            Fax: 81-298-36-9060
            Email: mseki@rtc.riken.go.jp
            An Arabidopsis full-length cDNA library was constructed essentially
            as reported previously (Seki et al., 1998).cDNA cleaved with BamHI
            and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
            al., submitted for publication) digested with BamHI and SalI. This
            clone is in a modified pBluescript vector. Please visit our web
            site (http://www.gsc.riken.go.jp/e/plant/index _e.html) for further
            details.
FEATURES             Location/Qualifiers
     source          1. .636
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="RAFL19-12-M11"
                     /clone_lib="RAFL19"
                     /tissue_type="mixture of silique and flower"
                     /lab_host="DH10B"
                     /note="Site_1: BamHI; Site_2: SalI; Subtraction Library"
BASE COUNT      151 a    167 c    158 g    159 t      1 others
ORIGIN

  Query Match             8.9%;  Score 77;  DB 9;  Length 636;
  Best Local Similarity   54.4%;  Pred. No. 2.1e-08;
  Matches  155;  Conservative    0;  Mismatches  130;  Indels    0;  Gaps    0;

Qy         92 AAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTAT 151
              | |||    ||  ||| |||||    || || || |||||  |  | ||  | || |
Db        318 AGGAAGCCGTGCGGTACTCATTGCTAGCGGGCGGAAAACGTGTGAGGCCTCTGCTCTGCA 377

Qy        152 TACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCAC 211
              |   |   ||      |||  | |      |  |||       ||    ||
Db        378 TTGCCGTCTGCGAGCTTGTGGGAGGCGACGAGGCTACTGCCATGTCAGCTGCTTGCGCGG 437

Qy        212 TAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATG 271
              | || ||||| || |||   || ||||||||||| || || ||    ||||| ||||  |
Db        438 TTGAGATGATCCACACAAGCTCTCTTATTCATGACGATCTTCCGTGCATGGACAATGCCG 497
```

```
Qy        272 ATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAG 331
              |    ||  |||| ||    || || || || |||||||| ||    |  ||| | || |
Db        498 ACCTCCGCAGAGGCAAGCCCACCAACCACAAGGTATATGGAGAAGACATGGCGGTTTTGG 557

Qy        332 CAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTG 376
              ||||||||||  |  |  |    || ||||| |  ||  |   ||
Db        558 CAGGTGATGCACTCCTTGCATTGGCGTTTGAGCACATGACGGTTG 602


RESULT 15
BQ506776
LOCUS       BQ506776                716 bp    mRNA    linear   EST 22-JUL-2002
DEFINITION  EST614191 Generation of a set of potato cDNA clones for microarray
            analyses mixed potato tissues Solanum tuberosum cDNA clone STMGM34
            5' end, mRNA sequence.
ACCESSION   BQ506776
VERSION     BQ506776.2  GI:21922629
KEYWORDS    EST.
SOURCE      potato.
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 716)
  AUTHORS   Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C.,
            Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and
            Karamycheva,S.A.
  TITLE     Generation of a set of potato cDNA clones for microarray analyses
  JOURNAL   Unpublished (2002)
COMMENT     On Jun 10, 2002 this sequence version replaced gi:21365645.
            Other_ESTs: EST614192
            Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato@tigr.org
            This clone is available through the Research Genetics, contact the
            Research Genetics for further information 1-800-711-6195 or
            cdna@resgen.com
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .716
                     /organism="Solanum tuberosum"
                     /cultivar="Kennebec or Binjte"
                     /db_xref="taxon:4113"
                     /clone="STMGM34"
                     /clone_lib="Generation of a set of potato cDNA clones for
                     microarray analyses mixed potato tissues"
                     /tissue_type="mixed tissues"
                     /lab_host="SOLR"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; supplier: Combination of untreated and Phytophthora
                     infestans-treated libraries of stolons, leaves, leaflets,
                     axillary buds of stem explants, petioles, germinating eyes
                     , tubers, or roots."
BASE COUNT      226 a    150 c    149 g    191 t
ORIGIN

  Query Match             8.8%;  Score 75.4;  DB 14;  Length 716;
  Best Local Similarity   53.6%;  Pred. No. 5.2e-08;
  Matches  157;  Conservative    0;  Mismatches  136;  Indels    0;  Gaps    0;

Qy         81 TACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCGACC 140
              ||   || |  | |||   |||   || ||| |   | ||| ||| || ||  |  | ||
Db        398 TATCAAGTTCCATGAAGCAATGAGATACTCACTCCTTTCTGAAGGCAAGCGGGTTTGCCC 457

Qy        141 AGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAG 200
               || || |   ||  | | |  ||     |   |  | | |     ||  |||
Db        458 TGTACTCTGCATAGCCGCCTGTGAGCTTGTTGGTGGCCAAGAATCAACAGCAATGCCTGC 517

Qy        201 CGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAGCGAT 260
               ||   || | | || ||||| ||| |    |   | || || || ||||| ||    ||
Db        518 TGCTTGTGGAATGGAGATGATGCATGCTATGTGGTTGATGCACGACGACCTTCCCTGCAT 577

Qy        261 GGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGAC 320
              ||| ||||||||||  || ||| |||  |  ||||||| || || |||||
Db        578 GGACAATGATGATTTCCGTCGAAGAAAGCTGTCAAATCACAAGGTTTATGGGCGAAACGT 637

Qy        321 TGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAA 373
               |  | ||||| ||| || |  |      |    ||||||  | | |||  |||
Db        638 CACTGTTTTAGCTGGTTATTCCCTTGGTGCCTTTGCATTTCAGCATATGACAA 690


Search completed: May 30, 2003, 17:23:48
Job time : 1605 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 29, 2003, 11:47:05 ; Search time 35 Seconds
(without alignments)
1092.655 Million cell updates/sec

Title: US-09-925-637-64
Perfect score: 1453
Sequence: 1 MTNLPMNKLIDEVNNELSVA.....ELTQIDEQFNTKHLLEIVDL 287

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1453	100.0	287	22	AAU00859	S. aureus Farnesyl
2	1442	99.2	293	22	AAU36990	Staphylococcus aur
3	1436	98.8	293	21	AAB23333	Staphylococcus aur
4	1415	97.4	288	22	AAU33743	Staphylococcus aur
5	1015	69.9	302	23	ABP38326	Staphylococcus epi
6	943	64.9	260	22	AAG81565	S. epidermidis ope
7	702.5	48.3	293	23	ABB48306	Listeria monocytog
8	657	45.2	297	17	AAW00285	Mutant farnesyldip
9	656	45.1	297	17	AAW00286	Native farnesyldip
10	654	45.0	297	19	AAW62532	Farnesyl diphospat
11	654	45.0	297	21	AAB29399	Bacillus stearothe
12	653	44.9	297	17	AAW00283	Mutant farnesyldip
13	650	44.7	297	17	AAW00284	Mutant farnesyldip
14	650	44.7	297	19	AAW47444	Bacillus stearothe
15	650	44.7	297	20	AAY27008	Farnesyl diphospha
16	647	44.5	297	14	AAR35047	FPS. Bacillus ste
17	646	44.5	297	20	AAY27006	Geranyl diphosphat
18	645	44.4	297	17	AAW00282	Mutant farnesyldip
19	644	44.3	297	20	AAY27007	Geranyl diphosphat
20	637	43.8	297	19	AAW62535	Mutant farnesyl di
21	637	43.8	297	19	AAW62537	Mutant farnesyl di
22	636	43.8	297	19	AAW62533	Mutant farnesyl di
23	618	42.5	295	19	AAW62536	Mutant farnesyl di
24	618	42.5	295	19	AAW62546	Mutant farnesyl di
25	617	42.5	295	19	AAW62534	Mutant farnesyl di
26	596	41.0	293	22	AAU35037	Enterococcus faeca
27	551.5	38.0	290	23	ABP25554	Streptococcus poly
28	551	37.9	291	21	AAB26267	Farnesyl diphospha
29	549	37.8	291	22	AAU38004	Streptococcus pneu
30	547	37.6	291	22	AAU37680	Streptococcus pneu
31	540.5	37.2	285	23	ABB54173	Lactococcus lactis
32	521	35.9	393	21	AAY82667	Taxus GGPP synthas
33	520.5	35.8	393	21	AAY82651	Taxus canadensis g
34	520.5	35.8	393	21	AAY82653	Taxus GGPP synthas
35	520.5	35.8	393	21	AAY82654	Taxus GGPP synthas
36	520.5	35.8	393	21	AAY82655	Taxus GGPP synthas
37	520.5	35.8	393	21	AAY82656	Taxus GGPP synthas
38	520.5	35.8	393	21	AAY82657	Taxus GGPP synthas
39	520.5	35.8	393	21	AAY82658	Taxus GGPP synthas
40	520.5	35.8	393	21	AAY82659	Taxus GGPP synthas
41	520.5	35.8	393	21	AAY82660	Taxus GGPP synthas
42	520.5	35.8	393	21	AAY82661	Taxus GGPP synthas
43	520.5	35.8	393	21	AAY82662	Taxus GGPP synthas
44	520.5	35.8	393	21	AAY82663	Taxus GGPP synthas
45	520.5	35.8	393	21	AAY82664	Taxus GGPP synthas

ALIGNMENTS

RESULT 1
AAU00859
ID AAU00859 standard; Protein; 287 AA.
XX
AC AAU00859;
XX
DT 04-JUL-2001 (first entry)
XX
DE S. aureus Farnesyl diphosphatesynthase, IspA.
XX
KW Farnesyl diphosphatesynthase; IspA; immunogen; vaccine; antibody;
KW wound infection; cellulitis; burn infection; eyelid infection;
KW food poisoning; joint infection; neonatal conjunctivitis; osteomyelitis;
KW skin infection; scalded skin syndrome; toxic epidermal necrosis;
KW Ritter's disease; Lyell's disease; toxic shock syndrome; endocarditis.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT Region 88..91
FT /label= Antigenic_epitope
FT Region 93..95
FT /label= Antigenic_epitope
FT Region 240..243
FT /label= Antigenic_epitope
XX
PN WO200116292-A2.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-US23773.
XX
PR 01-SEP-1999; 99US-0151933.

XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Choi GH;
XX
DR WPI; 2001-183259/18.
DR N-PSDB; AAS00821.
XX
PT New isolated nucleic acid for use in diagnosing Staphylococcus
PT infections and in vaccines for eliciting immune responses to the
PT infections -
XX
PS Claim 9; Page 23; 225pp; English.
XX
CC The sequence represents S. aureus IspA (Farnesyl diphosphatesynthase).
CC The polynucleotides of the invention are used to detect Staphylococcus
CC nucleic acids in a biological sample from an animal for diagnosing
CC Staphylococcus infections. The polypeptides of the invention are used to
CC detect anti-Staphylococcus antibodies in a biological sample from an
animal to diagnose Staphylococcus infections. The polypeptides are also
used in vaccines to elicit protective antibodies in an animal to a member
of the Staphylococcus genus and for preventing or attenuating an
CC infection caused by a member of the Staphylococcus genus e.g wound
CC infection, cellulitis, burn infection, eyelid infection, food poisoning,
CC joint infection, neonatal conjunctivitis, osteomyelitis, skin infection,
CC scalded skin syndrome (also known as toxic epidermal necrosis, Ritter's
CC disease and Lyell's disease), toxic shock syndrome and endocarditis. The
CC polynucleotides may also be used in vaccines and for preventing or
CC attenuating a Staphylococcus infection. Antibodies to the polypeptides
CC may be used to purify, detect and target the polypeptides in in vitro and
CC in vivo diagnostic and therapeutic methods.
XX
SQ Sequence 287 AA;

Query Match 100.0%; Score 1453; DB 22; Length 287;
Best Local Similarity 100.0%; Pred. No. 4.2e-125;
Matches 287; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MTNLPMNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTNLPMNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTE 60

Qy         61 YELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAF 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAF 120

Qy        121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGA 180

Qy        181 LLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLEN 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLEN 240

Qy        241 NKSTYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NKSTYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287
```

RESULT 2
AAU36990
ID AAU36990 standard; Protein; 293 AA.
XX
AC AAU36990;
XX
DT 14-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #1160.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.

XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS54849.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Example 3; Seq ID No 12583; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 293 AA;

Query Match 99.2%; Score 1442; DB 22; Length 293;
Best Local Similarity 99.3%; Pred. No. 4.4e-124;
Matches 285; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 MTNLPMNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTNLPMNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTE 60

Qy         61 YELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAF 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAF 120

Qy        121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGA 180

Qy        181 LLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLEN 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        181 LLTFAVMSAADIANVDDATKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLEN 240

Qy        241 NKSTYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287
              ||||||||||||||||||||||||||| ||||||||||||||||||
Db        241 NKSTYVSLLGKDGAEDKLTYHRDAAVDELMQIDEQFNTKHLLEIVDL 287
```

```
RESULT 3
AAB23333
ID   AAB23333 standard; Protein; 293 AA.
XX
AC   AAB23333;
XX
DT   12-JAN-2001  (first entry)
XX
DE   Staphylococcus aureus IspA.
XX
KW   IspA; bacterial disease; respiratory tract infection;
KW   gastrointestinal infection; cardiac infection; Helicobacter pylori;
KW   stomach cancer; stomach ulcer; gastritis.
XX
OS   Staphylococcus aureus.
XX
PN   US6107058-A.
XX
PD   22-AUG-2000.
XX
PF   26-MAR-1999;   99US-0276873.
XX
PR   26-MAR-1999;   99US-0276873.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI   Gwynn M,  Wilding EI;
XX
DR   WPI; 2000-578535/54.
DR   N-PSDB; AAA92031.
XX
PT   Novel farnesyl diphosphate synthase polynucleotide from staphylococcus
PT   aureus useful for diagnosis and treatment of bacterial infections and
PT   as hybridization probe for isolating genomic clones -
XX
PS   Claim 13; column 3-4; 15pp; English.
XX
CC   The present sequence is the Staphylococcus aureus IspA protein. This
CC   protein and its coding sequence can be used in many research assays, as
CC   well as treatments for bacterial diseases such as infections of the
CC   respiratory tract (including otitis media, bacterial tracheitis, acute
CC   epiglottitis, thyroiditis, empyema and lung abscesses), cardiac
CC   infections such as infective endocarditis, gastrointestinal infections
CC   including secretory diarrhoea, splenic abscesses and retroperitoneal
CC   abscesses, CNS infections such as cerebral abscesses, eye infections
CC   (including blepharitis, conjunctivitis, keratitis, endophthalmitis,
CC   preseptal and orbital cellulitis and darcryocystitis), kidney and urinary
CC   tract infections such as epididymitis, intrarenal and perinephric
CC   abscesses and toxic shock syndrome, skin diseases (including impetigo,
CC   folliculitis, cutaneous abscesses, wound infection and bacterial
CC   myositis), bone and joint infections such as septic arthritis and
CC   osteomyelitis, septic thrombophlebitis, food poisoning and scalded skin
CC   syndrome. In addition, they can be used to treat diseases caused by
CC   Helicobacter pylori, including stomach cancer, stomach ulcers and
CC   gastritis.
XX
SQ   Sequence   293 AA;

  Query Match             98.8%;  Score 1436;  DB 21;  Length 293;
  Best Local Similarity   99.0%;  Pred. No. 1.6e-123;
  Matches  284;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MTNLPMNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTNLPMNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTE 60

Qy         61 YELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAF 120
              ||||:|||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         61 YELGVKSAIALEMIHTYSLIHDDLPAMDNDDYRRRKLTNHKVYGEWTAILAGDALLTKAF 120

Qy        121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGA 180

Qy        181 LLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLEN 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        181 LLTFAVMSAADIANVDDATKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLEN 240

Qy        241 NKSTYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NKSTYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287


RESULT 4
AAU33743
ID   AAU33743 standard; Protein; 288 AA.
XX
AC   AAU33743;
XX
DT   14-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus cellular proliferation protein #19.
XX
KW   Antisense; prokaryotic cellular proliferation protein;
KW   antibiotic; antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   N-PSDB; AAS51602.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Example 3; Seq ID No 5239; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence represents an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
```

```
XX
SQ   Sequence   288 AA;

  Query Match            97.4%;  Score 1415;  DB 22;  Length 288;
  Best Local Similarity  99.3%;  Pred. No. 1.3e-121;
  Matches  280;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          6 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGM 65
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGM 60

Qy         66 KSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISS 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISS 120

Qy        126 DDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFA 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFA 180

          186 VMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTY 245
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
          181 VMSAADIANVDDATKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTY 240

Qy        246 VSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287
              ||||||||||||||||||||| ||||||||||||||||||||
Db        241 VSLLGKDGAEDKLTYHRDAAVDELMQIDEQFNTKHLLEIVDL 282


RESULT 5
ABP38326
ID   ABP38326 standard; Protein; 302 AA.
XX
AC   ABP38326;
XX
DT   24-JUL-2002  (first entry)
XX
DE   Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3171.
XX
KW   Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW   antibacterial; gene therapy.
XX
OS   Staphylococcus epidermidis.
XX
PN   US6380370-B1.
XX
PD   30-APR-2002.
XX
PF   13-AUG-1998;   98US-0134001.
XX
PR   14-AUG-1997;   97US-055779P.
PR   08-NOV-1997;   97US-064964P.
XX
PA   (GENO-) GENOME THERAPEUTICS CORP.
XX
PI   Doucette-Stamm LA,  Bush D;
XX
DR   WPI; 2002-381255/41.
DR   N-PSDB; ABN90871.
XX
PT   Novel isolated nucleic acid encoding a Staphylococcus epidermidis
PT   polypeptide, useful for diagnosing and treating bacterial infections -
XX
PS   Disclosure; SEQ ID 3171; 267pp; English.
XX
CC   ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC   frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC   given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC   antibacterial activity and can be used in gene therapy. The sequences
CC   can also be used in the diagnosis and treatment of bacterial infections,
CC   particularly S. epidermidis infections. The sequences can be used to
CC   screen for compounds able to interfere with the S. epidermidis life
CC   cycle or inhibit S. epidermidis infection.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the
CC   USPTO web site.
XX
SQ   Sequence   302 AA;

  Query Match            69.9%;  Score 1015;  DB 23;  Length 302;
  Best Local Similarity  68.6%;  Pred. No. 6.8e-85;
  Matches  197;  Conservative   41;  Mismatches   49;  Indels    0;  Gaps    0;

Qy          1 MTNLPMNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTE 60
              |  | |||||: :|  |: :|  | : | ||||| |||||||||||||:||||| || :
Db         10 MKKLQMNKLINIINTSLNKSIQSSPLKTNLEESMKYSLNAGGKRIRPVILLLTLKMLNKD 69

Qy         61 YELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAF 120
              |: |: ||:|||||||||||||||||||||||||||||||||| ||||||||||||
Db         70 YQQGLNSALALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWKAILAGDALLTKAF 129

Qy        121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGA 180
              ||:|:|  : | ||: :::||| ||||:|||||| |||:|||: | ||||| ||:|||||
Db        130 ELVSNDTTIEDSVKVSIIKRLSKASGHLGMVGGQALDMESEGKSIRLETLESIHETKTGA 189

Qy        181 LLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLEN 240
              || |:||:| ||| |:    ::|: :|:|||||||||||| ||||:|||||||: |
Db        190 LLNFSVMAAVDIAQVEQNIAKNLDEFSHHLGMMFQIKDDLLDVYGDESKLGKKVGSDIVN 249

Qy        241 NKSTYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287
              :|||||||||:|||:||  |:  |:: | || :|::|  | :||||
Db        250 HKSTYVSLLGKEGAEEKLNNHQYLAMNCLNQISDQYDTSELSDIVDL 296


RESULT 6
AAG81565
ID   AAG81565 standard; Protein; 260 AA.
XX
AC   AAG81565;
XX
DT   03-SEP-2001  (first entry)
XX
DE   S. epidermidis open reading frame protein sequence SEQ ID NO:224.
XX
KW   Staphylococcus epidermidis SR1 strain; infection; diagnosis;
KW   vaccination; endocarditis.
XX
OS   Staphylococcus epidermidis.
XX
PN   WO200134809-A2.
XX
PD   17-MAY-2001.
XX
PF   09-NOV-2000; 2000WO-US30782.
XX
PR   09-NOV-1999;   99US-0164258.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Kimmerly WJ;
XX
DR   WPI; 2001-316495/33.
DR   N-PSDB; AAH52415.
XX
PT   Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT   useful for vaccinating against infections, e.g. endocarditis -
XX
PS   Claim 18; Page 103; 2188pp; English.
XX
CC   AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC   (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC   (I) and (II) can have antibacterial activity and therefore can be used
CC   in vaccination. The nucleic acids (I) may be used to produce the
CC   S. epidermidis polypeptides (II) via the production of vectors
CC   containing them which are used to produce hosts cells which express the
CC   polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC   used to vaccinate subjects and to raise antibodies against the bacteria.
```

```
CC   The polypeptides may also be used to assay for other inhibitors of their
CC   activity and therefore identify compounds that may be used for the
CC   treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC   AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC   polynucleotide sequences from the present invention. AAH55091 to
CC   AAH55098 represent oligonucleotide sequences and primers which are used
CC   in the exemplification of the present invention.
CC   N.B. The present invention specifically claims all the polynucleotide
CC   sequences given in the sequence listing of the present specification,
CC   however the sequence listing only goes up to SEQ ID NO:4454 so even
CC   though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC   no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ   Sequence   260 AA;

  Query Match            64.9%;  Score 943;  DB 22;  Length 260;
  Best Local Similarity  71.3%;  Pred. No. 2.3e-78;
  Matches  181;  Conservative   36;  Mismatches   37;  Indels    0;  Gaps    0;

Qy         34 MLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYR 93
              | ||||||||||||||:|||||  || :|: |: ||:||||||||||||||||||||||||
Db          1 MKYSLNAGGKRIRPVILLLTLKMLNKDYQQGLNSALALEMIHTYSLIHDDLPAMDNDDYR 60

Qy         94 RGKLTNHKVYGEWTAILAGDALLTKAFELISSDDRLTDEVKIKVLQRLSIASGHVGMVGG 153
              |||||||||||| ||||||||||||||:|:|  : | ||: :::||| ||||:|||||
Db         61 RGKLTNHKVYGEWKAILAGDALLTKAFELVSNDTTIEDSVKVSIIKRLSKASGHLGMVGG 120

Qy        154 QMLDMQSEGQPIDLETLEMIHKTKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMM 213
              | |||:|||: | ||||| ||:|||||| |:||:| ||| |:    ::|: :|:|||||
Db        121 QALDMESEGKSIRLETLESIHETKTGALLNFSVMAAVDIAQVEQNIAKNLDEFSHHLGMM 180

Qy        214 FQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVSLLGKDGAEDKLTYHRDAAVDELTQID 273
              ||||||||| ||||:|||||||||: |:|||||||||||:|||:||  |:  |:: | ||
Db        181 FQIKDDLLDVYGDESKLGKKVGSDIVNHKSTYVSLLGKEGAEEKLNNHQYLAMNCLNQIS 240

Qy        274 EQFNTKHLLEIVDL 287
              :|:::|  | :||||
Db        241 DQYDTSELSDIVDL 254


RESULT 7
ABB48306
ID   ABB48306 standard; Protein; 293 AA.
XX
AC   ABB48306;
XX
DT   05-FEB-2002  (first entry)
XX
DE   Listeria monocytogenes protein #1010.

KW   Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW   vitamin B12; bacterial infection; disease.
XX
OS   Listeria monocytogenes.
XX
PN   WO200177335-A2.
XX
PD   18-OCT-2001.
XX
PF   11-APR-2001; 2001WO-FR01118.
XX
PR   11-APR-2000; 2000FR-0004629.
XX
PA   (INSP ) INST PASTEUR.
XX
PI   Buchrieser C,  Frangeul L,  Couve E,  Rusniok C,  Fsihi H,  Dehoux P;
PI   Dussurget O,  Chetouani F,  Nedjari H,  Glaser P,  Kunst F,  Cossart P;
PI   Daniels J,  Goebel W,  Kreft J,  Kuhn M,  Ng E,  Vazquez-Boland JA;
PI   Dominguez-Bernal G,  Garrido-Garcia P,  Tierrez-Martinez A,  Amend A;
PI   Chakraborty T,  Domann E,  Hain T,  Berche P,  Charbit A,  Durant L;
PI   Perez-Diaz J,  Baquero F,  Garcia Del Portillo F,  Gomez-Lopez N;
PI   Maduenio E,  De Pablos B,  Wehland J,  Kaerst U,  Entian K,  Hauf J;
PI   Rose M,  Voss H;
XX
DR   WPI; 2002-010914/01.
XX
PT   Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT   and prevention of Listeria and related bacterial infections, and
PT   related polypeptides -
XX
PS   Claim 6; SEQ ID No 1011; 192pp; French.
XX
CC   The present invention relates to the genome sequence of Listeria
CC   monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
CC   it are useful for selecting probes and primers for detecting genes in L.
CC   monocytogenes and related organisms, and for studying genetic
CC   polymorphisms and other genomes. The present sequence is a protein
CC   encoded by the genome sequence of the present invention. Proteins
CC   expressed from the genome sequence are useful for raising specific
CC   antibodies, identification of L. monocytogenes and related organisms, and
CC   for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC   B12. The genome sequence and proteins encoded by it are also useful for
CC   selecting compounds that regulate gene expression and cell replication
CC   and modulate L. monocytogenes-related diseases. In addition, the genome
CC   sequence and proteins encoded by it are useful in pharmaceutical and
CC   vaccines compositions for the treatment or prevention of infections by L.
CC   monocytogenes and related organisms.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   293 AA;

  Query Match            48.3%;  Score 702.5;  DB 23;  Length 293;
  Best Local Similarity  52.9%;  Pred. No. 3.1e-56;
  Matches  148;  Conservative   46;  Mismatches   81;  Indels    5;  Gaps    2;

Qy          8 KLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKS 67
              |:|||    |   ||:  :: :|:||||||: ||||||||:|:  || :|     ||:|:
Db         13 KVIDE---SLFKEINERNIEPRLKESMLYSIQAGGKRIRPMLVFATLQALKVNPLLGVKT 69

Qy         68 AIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSDD 127
              | |||||||||||||||||||||||||| |||||:|: ||||||||||| || ::: ||
Db         70 ATALEMIHTYSLIHDDLPAMDNDDYRRGKWTNHKVFGDATAILAGDALLTLAFSILAEDD 129

Qy        128 RLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFAVM 187
              |: | :| :: ::| :|| ||||||: |:::| : : || |  ||  ||| || :||
Db        130 NLSFETRIALINQISFSSGAEGMVGGQLADLEAENKQVTLEELSSIHARKTGELLIYAVT 189

Qy        188 SAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVS 247
              ||| ||  |    : |  :: ::|: ||| ||:||  ||| |:||| |:|   |||||
Db        190 SAAKIAEADPEQTKRLRIFAENIGIGFQISDDILDVIGDETKMGKKTGADAFLNKSTYPG 249

Qy        248 LLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287
              || |||:  |  |   |   |:  |  |: : ||:: ||
Db        250 LLTLDGAKRALNEHVTIAKSALSGHD--FDDEILLKLADL 287


RESULT 8
AAW00285
ID   AAW00285 standard; Protein; 297 AA.
XX
AC   AAW00285;
XX
DT   11-MAY-1997  (first entry)
XX
DE   Mutant farnesyldiphosphate synthase (4).
XX
KW   Farnesyldiphosphate; FPP; synthase; mutant; enzyme;
KW   geranylgeranyl diphosphate; GGPP.
XX
OS   Bacillus stearothermophilus.
XX
PN   EP733709-A2.
XX
PD   25-SEP-1996.
```

```
XX
PF   29-SEP-1995;   95EP-0115423.
XX
PR   14-FEB-1995;   95JP-0025253.
XX
PA   (TOYT ) TOYOTA JIDOSHA KK.
XX
PI   Ayumi K,  Kyozo O,  Shinichi O,  Shusei O,  Takeshi N;
PI   Tanetoshi K,  Tokuzo N;
XX
DR   WPI; 1996-427057/43.
DR   N-PSDB; AAT40228.
XX
PT   Production of geranyl:geranyl di:phosphate synthase enzymes - by
PT   mutation of farnesyl di:phosphate synthase
XX
PS   Claim 12; Page 15-16; 50pp; English.
XX
CC   Four mutant FPP synthase genes are given in AAT40225 to AAT40228.  The
     native FPP synthase gene is given in AAT40229.  The mutants are
     capable of synthesising geranylgeranyl diphosphate (GGPP) or
     geranylgeranyol from isopentenyl, dimethylallyl, geranyl or farnesyl
CC   diphosphate.  The GGPP synthase is produced from FPP synthase by
CC   deletion, addition or replacement of one to a few amino acids.  The
CC   modification is present in at least one of the following positions:
CC   34, 59, 81, 157, 182, 239, 265 and 275.
XX
SQ   Sequence   297 AA;

  Query Match             45.2%;  Score 657;  DB 17;  Length 297;
  Best Local Similarity   47.7%;  Pred. No. 4.7e-52;
  Matches  133;  Conservative   57;  Mismatches   83;  Indels    6;  Gaps    2;

Qy          1 MTNLPMNKLIDEVNNELSVAINKSVM----DTQLEESMLYSLNAGGKRIRPVLLLLTLDS 56
              |  | : : ::|    :  |::: :       :|:::| ||| ||||||||:||| |: :
Db          1 MAQLSVEQFLNEQKQAVETALSRYIERLEGPAKLKKAMAYSLEAGGKRIRPLLLLSTVRA 60

Qy         57 LNTEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALL 116
              |  :  :|:  | |:|||||:||||||||:||||| |||| |||||:|| |||||| ||
Db         61 LGKDPAVGLPVACAIEMIHTHSLIHDDLPSMDNDDLRRGKPTNHKVFGEAMAILAGDGLL 120

Qy        117 TKAFELIS--SDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIH 174
              | ||:||:   |:|:   |:::::::||: |:|  ||| ||  ||: ||: : |  || ||
Db        121 TYAFQLITEIDDERIPPSVRLRLIERLAKAAGPEGMVAGQAADMEGEGKTLTLSELEYIH 180

Qy        175 KTKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKV 234
              : ||| :| ::| : | |   |      |: :: |||: |||:||:||  | | |:||:|
Db        181 RHKTGKMLQYSVHAGALIGGADARQTRELDEFAAHLGLAFQIRDDILDIEGAEEKIGKRV 240

          235 GSDLENNKSTYVSLLGKDGAEDKLTYHRDAAVDELTQID 273
              |||  |||:|| :||   ||::|||:| :||   |   |
Db        241 GSDQSNNKATYPALLSLAGAKEKLTFHIEAAQRHLRNAD 279


RESULT 9
AAW00286
ID   AAW00286 standard; Protein; 297 AA.
XX
AC   AAW00286;
XX
DT   11-MAY-1997  (first entry)
XX
DE   Native farnesyldiphosphate synthase.
XX
KW   Farnesyldiphosphate; FPP; synthase; mutant; enzyme;
KW   geranylgeranyl diphosphate; GGPP.
XX
OS   Bacillus stearothermophilus.
XX
PN   EP733709-A2.
XX
PD   25-SEP-1996.
XX
PF   29-SEP-1995;   95EP-0115423.
XX
PR   14-FEB-1995;   95JP-0025253.
XX
PA   (TOYT ) TOYOTA JIDOSHA KK.
XX
PI   Ayumi K,  Kyozo O,  Shinichi O,  Shusei O,  Takeshi N;
PI   Tanetoshi K,  Tokuzo N;
XX
DR   WPI; 1996-427057/43.
DR   N-PSDB; AAT40229.
XX
PT   Production of geranyl:geranyl di:phosphate synthase enzymes - by
PT   mutation of farnesyl di:phosphate synthase
XX
PS   Disclosure; Page 17-19; 50pp; English.
XX
CC   Four mutant FPP synthase genes are given in AAT40225 to AAT40228.  The
CC   native FPP synthase gene is given in AAT40229.  The mutants are
CC   capable of synthesising geranylgeranyl diphosphate (GGPP) or
CC   geranylgeranyol from isopentenyl, dimethylallyl, geranyl or farnesyl
CC   diphosphate.  The GGPP synthase is produced from FPP synthase by
CC   deletion, addition or replacement of one to a few amino acids.  The
CC   modification is present in at least one of the following positions:
CC   34, 59, 81, 157, 182, 239, 265 and 275.
XX
SQ   Sequence   297 AA;

  Query Match             45.1%;  Score 656;  DB 17;  Length 297;
  Best Local Similarity   47.7%;  Pred. No. 5.8e-52;
  Matches  133;  Conservative   55;  Mismatches   85;  Indels    6;  Gaps    2;

Qy          1 MTNLPMNKLIDEVNNELSVAINKSVM----DTQLEESMLYSLNAGGKRIRPVLLLLTLDS 56
              |  | : : ::|    :  |::: :       :|:::| ||| ||||||||:||| |: :
Db          1 MAQLSVEQFLNEQKQAVETALSRYIERLEGPAKLKKAMAYSLEAGGKRIRPLLLLSTVQA 60

Qy         57 LNTEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALL 116
              |  :  :|:  | |:|||||||||||||:||||| |||| |||||:|| |||||| ||
Db         61 LGKDPAVGLPVACAIEMIHTYSLIHDDLPSMDNDDLRRGKPTNHKVFGEAMAILAGDGLL 120

Qy        117 TKAFELIS--SDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIH 174
              | ||:||:   |:|:   |::::::||: |:|  ||| ||  ||: ||: : |  || ||
Db        121 TYAFQLITEIDDERIPPSVRLRLIERLAKAAGPEGMVAGQAADMEGEGKTLTLSELEYIH 180

Qy        175 KTKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKV 234
              : ||| :| ::| : | |   |      |: :: |||: |||:||:||  | | |:|| |
Db        181 RHKTGKMLQYSVHAGALIGGADARQTRELDEFAAHLGLAFQIRDDILDIEGAEEKIGKPV 240

Qy        235 GSDLENNKSTYVSLLGKDGAEDKLTYHRDAAVDELTQID 273
              |||  |||:|| :||   ||::|| :| :||   |   |
Db        241 GSDQSNNKATYPALLSLAGAKEKLAFHIEAAQRHLRNAD 279


RESULT 10
AAW62532
ID   AAW62532 standard; Protein; 297 AA.
XX
AC   AAW62532;
XX
DT   14-SEP-1998  (first entry)
XX
DE   Farnesyl diphospate synthase of B. stearothermophilus.
XX
KW   Farnesyl diphosphate synthase; enzyme; prenyl diphosphate synthase;
KW   PDPS enzyme.
XX
OS   Bacillus stearothermophilus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "encoded by GTG"
FT   Domain          86..92
FT                   /note= "Asp-rich domain"
```

XX
PN WO9820138-A1.
XX
PD 14-MAY-1998.
XX
PF 29-OCT-1997; 97WO-JP03921.
XX
PR 05-NOV-1996; 96JP-0307506.
XX
PA (TOYT) TOYOTA JIDOSHA KK.
XX
PI Hirooka K, Nakane H, Nishino T, Ohnuma S, Ohto C;
XX
DR WPI; 1998-286954/25.
DR N-PSDB; AAV38455.
XX
PT New mutated prenyl diphosphate synthases and related nucleic acid,
PT and vectors - for synthesis of prenyl diphosphates longer than those
PT produced by wild-type enzyme
XX
PS Disclosure; Pages 31-32; 47pp; English.
XX
CC The present sequence represents a farnesyl diphosphate synthase
CC enzyme of Bacillus stearothermophilus. The enzyme has an Asp-rich
CC domains. The amino acid 8 and 5 positions upstream of the N-terminus of
CC the Asp-rich domain is replaced to produce a mutant prenyl diphosphate
CC synthase (PDPS) enzymes (AAW62533-37, AAW62546). The mutants can produce
CC longer chain products than wild-type enzymes. Mutant PDPS enzymes are
CC used to produce prenyl diphosphates of at least 20 carbon atoms, which
CC are precursors for steroids, ubiquinones, dolichols, carotenoids,
CC prenylated proteins, animal and plant hormones etc.
XX
SQ Sequence 297 AA;

Query Match 45.0%; Score 654; DB 19; Length 297;
Best Local Similarity 47.7%; Pred. No. 8.8e-52;
Matches 133; Conservative 55; Mismatches 85; Indels 6; Gaps 2;

```
Qy          1 MTNLPMNKLIDEVNNELSVAINKSVM----DTQLEESMLYSLNAGGKRIRPVLLLLTLDS 56
              |  | : : ::|    :  |::: :       :|:::| ||| |||||||||:||| |: :
Db          1 MAQLSVEQFLNEQKQAVETALSRYIERLEGPAKLKKAMAYSLEAGGKRIRPLLLLSTVRA 60

Qy         57 LNTEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALL 116
              |  :  :|:  | |:||||||||||||||:|||| |||| |||||:||  |||||| ||
Db         61 LGKDPAVGLPVACAIEMIHTYSLIHDDLPSMDNDDLRRGKPTNHKVFGEAMAILAGDGLL 120

Qy        117 TKAFELIS--SDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIH 174
              | ||:||:   |:|:   |:::::::||: |:|  ||| ||  ||: ||: : |  || ||
Db        121 TYAFQLITEIDDERIPPSVRLRLIERLAKAAGPEGMVAGQAADMEGEGKTLTLSELEYIH 180

Qy        175 KTKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKV 234
              : ||| :| ::| : | |    |      |: :: |||: |||:||:||  | | |:|| |
Db        181 RHKTGKMLQYSVHAGALIGGADARQTRELDEFAAHLGLAFQIRDDILDIEGAEEKIGKPV 240

Qy        235 GSDLENNKSTYVSLLGKDGAEDKLTYHRDAAVDELTQID 273
              |||  |||:|| :||   ||::|| :| :||   |   |
Db        241 GSDQSNNKATYPALLSLAGAKEKLAFHIEAAQRHLRNAD 279
```

RESULT 11
AAB29399
ID AAB29399 standard; Protein; 297 AA.
XX
AC AAB29399;
XX
DT 09-FEB-2001 (first entry)
XX
DE Bacillus stearothermophilus farnesyl diphosphate synthase.
XX
KW Farnesyl diphosphate synthase; transgenic microorganism;
KW monoterpene production; geranyl diphosphate synthase activity;
KW monoterpene synthase.
XX
OS Bacillus stearothermophilus.
XX
PN JP2000245482-A.
XX
PD 12-SEP-2000.
XX
PF 05-MAR-1999; 99JP-0059431.
XX
PR 05-MAR-1999; 99JP-0059431.
XX
PA (SOZO-) SOZOTEKI SEIBUTSU KOGAKU KENKYUSHO KK.
XX
DR WPI; 2000-642026/62.
DR N-PSDB; AAC63809.
XX
PT New transformed microbe useful for the production of monoterpene -
XX
PS Claim 5; Page 9-10; 15pp; Japanese.
XX
CC The invention relates to a transformed microorgansim containing a first
CC gene encoding a protein with geranyl diphosphate synthase activity,
CC and a second gene encoding a monoterpene synthase. The microorganism
CC is able to produce the particular monoterpene whose synthesis is
CC catalysed by the monoterpene synthase that has been introduced. The
CC invention also relates to the production of monoterpenes using the
CC novel microorganism. The microorganism can be used for the production of
CC various monoterpenes which are useful in industry. The present
CC sequence represents Bacillus stearothermophilus farnesyl disphosphate
CC synthase, a gene encoding which is claimed for use in a microorganism
CC of the invention.
XX
SQ Sequence 297 AA;

Query Match 45.0%; Score 654; DB 21; Length 297;
Best Local Similarity 47.7%; Pred. No. 8.8e-52;
Matches 133; Conservative 55; Mismatches 85; Indels 6; Gaps 2;

```
Qy          1 MTNLPMNKLIDEVNNELSVAINKSVM----DTQLEESMLYSLNAGGKRIRPVLLLLTLDS 56
              |  | : : ::|    :  |::: :       :|:::| ||| |||||||||:||| |: :
Db          1 MAQLSVEQFLNEQKQAVETALSRYIERLEGPAKLKKAMAYSLEAGGKRIRPLLLLSTVRA 60

Qy         57 LNTEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALL 116
              |  :  :|:  | |:||||||||||||||:|||| |||| |||||:||  |||||| ||
Db         61 LGKDPAVGLPVACAIEMIHTYSLIHDDLPSMDNDDLRRGKPTNHKVFGEAMAILAGDGLL 120

Qy        117 TKAFELIS--SDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIH 174
              | ||:||:   |:|:   |:::::::||: |:|  ||| ||  ||: ||: : |  || ||
Db        121 TYAFQLITEIDDERIPPSVRLRLIERLAKAAGPEGMVAGQAADMEGEGKTLTLSELEYIH 180

Qy        175 KTKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKV 234
              : ||| :| ::| : | |    |      |: :: |||: |||:||:||  | | |:|| |
Db        181 RHKTGKMLQYSVHAGALIGGADARQTRELDEFAAHLGLAFQIRDDILDIEGAEEKIGKPV 240

Qy        235 GSDLENNKSTYVSLLGKDGAEDKLTYHRDAAVDELTQID 273
              |||  |||:|| :||   ||::|| :| :||   |   |
Db        241 GSDQSNNKATYPALLSLAGAKEKLAFHIEAAQRHLRNAD 279
```

RESULT 12
AAW00283
ID AAW00283 standard; Protein; 297 AA.
XX
AC AAW00283;
XX
DT 11-MAY-1997 (first entry)
XX
DE Mutant farnesyldiphosphate synthase (2).
XX
KW Farnesyldiphosphate; FPP; synthase; mutant; enzyme;
KW geranylgeranyl diphosphate; GGPP.
XX
OS Bacillus stearothermophilus.
XX

PN EP733709-A2.
XX
PD 25-SEP-1996.
XX
PF 29-SEP-1995; 95EP-0115423.
XX
PR 14-FEB-1995; 95JP-0025253.
XX
PA (TOYT) TOYOTA JIDOSHA KK.
XX
PI Ayumi K, Kyozo O, Shinichi O, Shusei O, Takeshi N;
PI Tanetoshi K, Tokuzo N;
XX
DR WPI; 1996-427057/43.
DR N-PSDB; AAT40226.
XX
PT Production of geranyl:geranyl di:phosphate synthase enzymes - by
PT mutation of farnesyl di:phosphate synthase
XX
PS Claim 10; Page 10-12; 50pp; English.
XX
CC Four mutant FPP synthase genes are given in AAT40225 to AAT40228. The
CC native FPP synthase gene is given in AAT40229. The mutants are
CC capable of synthesising geranylgeranyl diphosphate (GGPP) or
CC geranylgeranyol from isopentenyl, dimethylallyl, geranyl or farnesyl
CC diphosphate. The GGPP synthase is produced from FPP synthase by
CC deletion, addition or replacement of one to a few amino acids. The
CC modification is present in at least one of the following positions:
CC 34, 59, 81, 157, 182, 239, 265 and 275.
XX
SQ Sequence 297 AA;

Query Match 44.9%; Score 653; DB 17; Length 297;
Best Local Similarity 47.3%; Pred. No. 1.1e-51;
Matches 132; Conservative 56; Mismatches 85; Indels 6; Gaps 2;

```
Qy          1 MTNLPMNKLIDEVNNELSVAINKSVM----DTQLEESMLYSLNAGGKRIRPVLLLLTLDS 56
              |  | : : ::|    :  |::: :      :::::| ||| ||||||||:||| |: :
Db          1 MAQLSVEQFLNEQKQAVETALSRYIERLEGPAKVKKAMAYSLEAGGKRIRPLLLLSTVQA 60

Qy         57 LNTEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALL 116
              |  :  :|:  | |:||||||||||||:|||| |||| |||||:||  |||||| ||
Db         61 LGKDPAVGLPVACAIEMIHTYSLIHDDLPSMDNDDLRRGKPTNHKVFGEAMAILAGDGLL 120

Qy        117 TKAFELIS--SDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIH 174
              | ||:||:   |:|:   |:::::::||: |:|  ||| ||  ||: ||: : |  || ||
Db        121 TYAFQLITEIDDERIPPSVRLRLIERLAKAAGPEGMVAGQAADMEGEGKTLTLSELEYIH 180

          175 KTKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKV 234
              : ||| :| ::| : | |   |      |: :: |||: |||:||:||  | | |:|| |
Db        181 RHKTGKMLQYSVHAGALIGGADARQTRELDEFAAHLGLAFQIRDDILDIEGAEEKIGKPV 240

Qy        235 GSDLENNKSTYVSLLGKDGAEDKLTYHRDAAVDELTQID 273
              |||  |||:|| :||   ||::|| :| :||   |   |
Db        241 GSDQSNNKATYPALLSLAGAKEKLAFHIEAAQRHLRNAD 279
```

RESULT 13
AAW00284
ID AAW00284 standard; Protein; 297 AA.
XX
AC AAW00284;
XX
DT 11-MAY-1997 (first entry)
XX
DE Mutant farnesyldiphosphate synthase (3).
XX
KW Farnesyldiphosphate; FPP; synthase; mutant; enzyme;
KW geranylgeranyl diphosphate; GGPP.
XX
OS Bacillus stearothermophilus.
XX
PN EP733709-A2.
XX
PD 25-SEP-1996.
XX
PF 29-SEP-1995; 95EP-0115423.
XX
PR 14-FEB-1995; 95JP-0025253.
XX
PA (TOYT) TOYOTA JIDOSHA KK.
XX
PI Ayumi K, Kyozo O, Shinichi O, Shusei O, Takeshi N;
PI Tanetoshi K, Tokuzo N;
XX
DR WPI; 1996-427057/43.
DR N-PSDB; AAT40227.
XX
PT Production of geranyl:geranyl di:phosphate synthase enzymes - by
PT mutation of farnesyl di:phosphate synthase
XX
PS Claim 11; Page 13-14; 50pp; English.
XX
CC Four mutant FPP synthase genes are given in AAT40225 to AAT40228. The
CC native FPP synthase gene is given in AAT40229. The mutants are
CC capable of synthesising geranylgeranyl diphosphate (GGPP) or
CC geranylgeranyol from isopentenyl, dimethylallyl, geranyl or farnesyl
CC diphosphate. The GGPP synthase is produced from FPP synthase by
CC deletion, addition or replacement of one to a few amino acids. The
CC modification is present in at least one of the following positions:
CC 34, 59, 81, 157, 182, 239, 265 and 275.
XX
SQ Sequence 297 AA;

Query Match 44.7%; Score 650; DB 17; Length 297;
Best Local Similarity 47.3%; Pred. No. 2e-51;
Matches 132; Conservative 55; Mismatches 86; Indels 6; Gaps 2;

```
Qy          1 MTNLPMNKLIDEVNNELSVAINKSVM----DTQLEESMLYSLNAGGKRIRPVLLLLTLDS 56
              |  | : : ::|    :  |::: :      :|:::| ||| ||||||||:||| |: :
Db          1 MAQLSVEQFLNEQKQAVETALSRYIERLEGPAKLKKAMAYSLEAGGKRIRPLLLLSTVRA 60

Qy         57 LNTEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALL 116
              |  :  :|:  | |:||||||||||||:|||| |||| |||||:||  |||||| ||
Db         61 LGKDPAVGLPVACAIEMIHTYSLIHDDLPSMDNDDLRRGKPTNHKVFGEAMAILAGDGLL 120

Qy        117 TKAFELIS--SDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIH 174
              | ||:||:   |:|:   |:::::::||: |:|  ||  ||  ||: ||: : |  || ||
Db        121 TYAFQLITEIDDERIPPSVRLRLIERLAKAAGPEGMAAGQAADMEGEGKTLTLSELEYIH 180

Qy        175 KTKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKV 234
              : ||| :| ::| : | |   |      |: :: |||: |||:||:||  | | |:|| |
Db        181 RYKTGKMLQYSVHAGALIGGADARQTRELDEFAAHLGLAFQIRDDILDIEGAEEKIGKPV 240

Qy        235 GSDLENNKSTYVSLLGKDGAEDKLTYHRDAAVDELTQID 273
              |||  |||:|| :||   ||::|| :| :||   |   |
Db        241 GSDQSNNKATYPALLSLAGAKEKLAFHIEAAQRHLRNAD 279
```

RESULT 14
AAW47444
ID AAW47444 standard; Protein; 297 AA.
XX
AC AAW47444;
XX
DT 05-JUN-1998 (first entry)
XX
DE Bacillus stearothermophilus farnesyl diphosphate synthase.
XX
KW Farnesyl diphosphate synthase; prenyl diphosphate synthase; mutant;
KW prenyl diphosphate.
XX
OS Bacillus stearothermophilus.
XX
PN EP816490-A2.
XX

```
PD   07-JAN-1998.
XX
PF   02-JUL-1997;   97EP-0111022.
XX
PR   03-JUL-1996;   96JP-0191635.
XX
PA   (TOYT ) TOYOTA JIDOSHA KK.
XX
PI   Ishida C,  Narita K,  Nishino T,  Ohnuma S,  Ohto C;
PI   Takeuchi Y;
XX
DR   WPI; 1998-054911/06.
DR   N-PSDB; AAV18516.
XX
PT   Mutant prenyl di:phosphate synthase - useful for producing
PT   long-chain prenyl di:phosphate
XX
PS   Claim 4; Page -; 25pp; English.
XX
CC   The present sequence is Bacillus stearothermophilus farnesyl
CC   diphosphate synthase, an example of a prenyl diphosphate synthase
CC   (PDS).
CC   A mutant PDS comprises an amino acid substitution located at the
CC   5th position in the amino-terminal direction from the 1st Asp of
CC   the N-terminal of the Asp-rich domain Asp-Asp-Xaa-Xaa-(Xaa-Xaa)-Asp
CC   (the Xaa in the parentheses may be absent), which is present in the
CC   2nd region among the conserved regions of wild type PDS. The mutant
CC   PDS can be used to produce a prenyl diphosphate having at least 20
CC   carbons from an isopentenyl diphosphate, dimethylallyl diphosphate,
CC   geranyl diphosphate, farnesyl diphosphate or geranylgeranyl
CC   diphosphate substrate. The substituted amino acid is involved in
CC   controlling the chain length of the prenyl diphosphate produced.
CC   N.B. Sequence not given in the specification, but decoded from the
CC   DNA sequence AAV18516.
XX
SQ   Sequence   297 AA;

  Query Match             44.7%;  Score 650;  DB 19;  Length 297;
  Best Local Similarity   49.6%;  Pred. No. 2e-51;
  Matches  134;  Conservative   49;  Mismatches   85;  Indels    2;  Gaps    1;

Qy          6 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGM 65
              :|:    |   ||  | :     :|:::| ||| |||||||||:||| |: :|  :  :|:
Db         10 LNEQKQAVETALSRYIERLEGPAKLKKAMAYSLEAGGKRIRPLLLLSTVRALGKDPAVGL 69

Qy         66 KSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELIS- 124
               | |:|||||||||||||:||||| |||| |||||:||  |||||| ||| ||:||:
Db         70 PVACAIEMIHTYSLIHDDLPSMDNDDLRRGKPTNHKVFGEAMAILAGDGLLTYAFQLITE 129

Qy        125 -SDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLT 183
               |:|:   |:::::::||: |:||  ||| ||  |+: ||: : +  || ||: ||| :|
Db        130 IDDERIPPSVRLRLIERLAKAAGPEGMVAGQAADMEGEGKTLTLSELEYIHRHKTGKMLQ 189

Qy        184 FAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKS 243
              .::| : | |   |      |: :: |||: |||:||:||  | | |:|| ||||  |||:
Db        190 YSVHAGALIGGADARQTRELDEFAAHLGLAFQIRDDILDIEGAEEKIGKPVGSDQSNNKA 249

Qy        244 TYVSLLGKDGAEDKLTYHRDAAVDELTQID 273
              || :||   ||::|| :| :||   |   |
Db        250 TYPALLSLAGAKEKLAFHIEAAQRHLRNAD 279


RESULT 15
AAY27008
ID   AAY27008 standard; Protein; 297 AA.
XX
AC   AAY27008;
XX
DT   20-SEP-1999  (first entry)
XX
DE   Farnesyl diphosphate synthase.
XX
KW   Geranyl diphosphate synthase; farnesyl diphosphate synthase; enzyme;
KW   Bacillus stearothermophilus; fermentation; geranyl monoterpene.
XX
OS   Bacillus stearothermophilus.
XX
PN   WO9931254-A1.
XX
PD   24-JUN-1999.
XX
PF   10-DEC-1998;   98WO-JP05590.
XX
PR   16-DEC-1997;   97JP-0346686.
XX
PA   (TOYT ) TOYOTA JIDOSHA KK.
XX
PI   Narita K,  Nishino T,  Ohnuma S,  Ohto C;
XX
DR   WPI; 1999-430040/36.
DR   N-PSDB; AAX86781.
XX
PT   Geranyl diphosphate synthase and gene encoding it
XX
PS   Disclosure; Page 35-37; 41pp; Japanese.
XX
CC   The invention provides a geranyl diphosphate synthase which is derived
CC   from the farnesyl diphosphate synthase of Bacillus stearothermophilus by
CC   replacement of the serine residue at position 82 with another residue of
CC   higher molecular weight. The enzyme is used in the fermentation and
CC   enzyme industries for efficient production of geranyl monoterpenes. The
CC   present sequence represents a farnesyl diphosphate synthase.
XX
SQ   Sequence   297 AA;

  Query Match             44.7%;  Score 650;  DB 20;  Length 297;
  Best Local Similarity   49.6%;  Pred. No. 2e-51;
  Matches  134;  Conservative   49;  Mismatches   85;  Indels    2;  Gaps    1;

Qy          6 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGM 65
              :|:    |   ||  | :     :|:::| ||| |||||||||:||| |: :|  :  :|:
Db         10 LNEQKQAVETALSRYIERLEGPAKLKKAMAYSLEAGGKRIRPLLLLSTVRALGKDPAVGL 69

Qy         66 KSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELIS- 124
               | |:|||||||||||||:||||| |||| |||||:||  |||||| ||| ||:||:
Db         70 PVACAIEMIHTYSLIHDDLPSMDNDDLRRGKPTNHKVFGEAMAILAGDGLLTYAFQLITE 129

Qy        125 -SDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLT 183
               |:|:   |:::::::||: |:||  ||| ||  ||: ||: : +  || ||: ||| :|
Db        130 IDDERIPPSVRLRLIERLAKAAGPEGMVAGQAADMEGEGKTLTLSELEYIHRHKTGKMLQ 189

Qy        184 FAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKS 243
              ::| : | |   |      |: :: |||: |||:||:||  | | |:|| ||||  |||:
Db        190 YSVHAGALIGGADARQTRELDEFAAHLGLAFQIRDDILDIEGAEEKIGKPVGSDQSNNKA 249

Qy        244 TYVSLLGKDGAEDKLTYHRDAAVDELTQID 273
              || :||   ||::|| :| :||   |   |
Db        250 TYPALLSLAGAKEKLAFHIEAAQRHLRNAD 279


Search completed: May 29, 2003, 11:58:54
Job time : 36 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 29, 2003, 11:58:20 ; Search time 15 Seconds
(without alignments)
562.959 Million cell updates/sec

Title: US-09-925-637-64
Perfect score: 1453
Sequence: 1 MTNLPMNKLIDEVNNELSVA.....ELTQIDEQFNTKHLLEIVDL 287

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1436	98.8	293	3	US-09-276-873-2	Sequence 2, Appli
2	1015	69.9	302	4	US-09-134-001C-3171	Sequence 3171, Ap
3	657	45.2	297	1	US-08-534-910B-9	Sequence 9, Appli
4	654	45.0	297	3	US-08-886-466-2	Sequence 2, Appli
5	654	45.0	297	4	US-09-475-304-2	Sequence 2, Appli
6	654	45.0	297	4	US-09-101-126-3	Sequence 3, Appli
7	653	44.9	297	1	US-08-534-910B-7	Sequence 7, Appli
8	650	44.7	297	1	US-08-534-910B-8	Sequence 8, Appli
9	650	44.7	297	1	US-08-534-910B-10	Sequence 10, Appl
10	650	44.7	297	4	US-09-367-528A-5	Sequence 5, Appli
11	646	44.5	297	4	US-09-367-528A-1	Sequence 1, Appli
12	645	44.4	297	1	US-08-534-910B-6	Sequence 6, Appli
13	644	44.3	297	4	US-09-367-528A-3	Sequence 3, Appli
14	551	37.9	291	4	US-09-275-742-2	Sequence 2, Appli
15	521	35.9	393	3	US-09-187-050-34	Sequence 34, Appl
16	520.5	35.8	393	3	US-09-187-050-2	Sequence 2, Appli
17	520.5	35.8	393	3	US-09-187-050-14	Sequence 14, Appl
18	520.5	35.8	393	3	US-09-187-050-16	Sequence 16, Appl
19	520.5	35.8	393	3	US-09-187-050-18	Sequence 18, Appl
20	520.5	35.8	393	3	US-09-187-050-20	Sequence 20, Appl
21	520.5	35.8	393	3	US-09-187-050-22	Sequence 22, Appl
22	520.5	35.8	393	3	US-09-187-050-24	Sequence 24, Appl
23	520.5	35.8	393	3	US-09-187-050-26	Sequence 26, Appl
24	520.5	35.8	393	3	US-09-187-050-27	Sequence 27, Appl
25	520.5	35.8	393	3	US-09-187-050-28	Sequence 28, Appl
26	520.5	35.8	393	3	US-09-187-050-29	Sequence 29, Appl
27	520.5	35.8	393	3	US-09-187-050-30	Sequence 30, Appl
28	520.5	35.8	393	3	US-09-187-050-31	Sequence 31, Appl
29	520.5	35.8	393	3	US-09-187-050-32	Sequence 32, Appl
30	520.5	35.8	393	3	US-09-187-050-33	Sequence 33, Appl
31	517	35.6	285	3	US-09-187-050-12	Sequence 12, Appl
32	465.5	32.0	377	4	US-09-420-211-2	Sequence 2, Appli
33	376.5	25.9	302	1	US-07-783-705A-1	Sequence 1, Appli
34	346	23.8	298	1	US-08-095-726-4	Sequence 4, Appli
35	346	23.8	298	1	US-08-096-043-4	Sequence 4, Appli
36	346	23.8	298	1	US-08-093-577-4	Sequence 4, Appli
37	346	23.8	298	1	US-08-096-623A-4	Sequence 4, Appli
38	346	23.8	307	1	US-08-095-726-2	Sequence 2, Appli
39	346	23.8	307	1	US-08-096-043-2	Sequence 2, Appli
40	346	23.8	307	1	US-08-093-577-2	Sequence 2, Appli
41	346	23.8	307	1	US-08-096-623A-2	Sequence 2, Appli
42	330	22.7	330	1	US-08-410-167A-4	Sequence 4, Appli
43	330	22.7	330	2	US-08-898-560-1	Sequence 1, Appli
44	323	22.2	330	4	US-09-101-126-1	Sequence 1, Appli
45	309.5	21.3	295	3	US-08-660-645A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-276-873-2
; Sequence 2, Application US/09276873
; Patent No. 6107058
; GENERAL INFORMATION:
; APPLICANT: Wilding, Edwina Imogen
; APPLICANT: Gwynn, Michael
; TITLE OF INVENTION: IspA
; FILE REFERENCE: GM10208
; CURRENT APPLICATION NUMBER: US/09/276,873
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-276-873-2

Query Match 98.8%; Score 1436; DB 3; Length 293;
Best Local Similarity 99.0%; Pred. No. 1e-132;
Matches 284; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 MTNLPMNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTNLPMNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTE 60

Qy         61 YELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAF 120
              ||||:|||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db         61 YELGVKSAIALEMIHTYSLIHDDLPAMDNDDYRRRKLTNHKVYGEWTAILAGDALLTKAF 120

Qy        121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGA 180

Qy        181 LLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLEN 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        181 LLTFAVMSAADIANVDDATKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLEN 240

Qy        241 NKSTYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NKSTYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287
```

RESULT 2
US-09-134-001C-3171
; Sequence 3171, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al

```
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3171
;   LENGTH: 302
;   TYPE: PRT
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3171

  Query Match             69.9%;  Score 1015;  DB 4;  Length 302;
  Best Local Similarity   68.6%;  Pred. No. 1.7e-91;
  Matches  197;  Conservative   41;  Mismatches   49;  Indels    0;  Gaps    0;

Qy          1 MTNLPMNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTE 60
              |  | |||||: :|  |: :|  | : | ||||| ||||||||||||:|||||  || :
Db         10 MKKLQMNKLINIIINTSLNKSIQSSPLKTNLEESMKYSLNAGGKRIRPVILLLTLKMLNKD 69

Qy         61 YELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAF 120
              |: |: ||:|||||||||||||||||||||||||||||||||||||| ||||||||||||
Db         70 YQQGLNSALALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWKAILAGDALLTKAF 129

Qy        121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGA 180
              ||:|:|  : | ||: :::||| ||||:|||||| |||:|||: | ||||| ||:|||||
Db        130 ELVSNDTTIEDSVKVSIIKRLSKASGHLGMVGGQALDMESEGKSIRLETLESIHETKTGA 189

Qy        181 LLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLEN 240
              || |:||:| ||| |:    ::|: :|:||||||||||||| ||||:|||||||: |
Db        190 LLNFSVMAAVDIAQVEQNIAKNLDEFSHHLGMMFQIKDDLLDVYGDESKLGKKVGSDIVN 249

Qy        241 NKSTYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287
              :|||||||||:|||:||  |:  |:: | || :|::|  | :||||
Db        250 HKSTYVSLLGKEGAEEKLNNHQYLAMNCLNQISDQYDTSELSDIVDL 296


RESULT 3
US-08-534-910B-9
; Sequence 9, Application US/08534910B
; Patent No. 5766911
;  GENERAL INFORMATION:
;    APPLICANT:  KOIKE, Ayumi
;    APPLICANT:  OBATA, Shusei
;    APPLICANT:  NISHINO, Tokuzo
;    APPLICANT:  OHNUMA, Shinichi
;    APPLICANT:  NAKAZAWA, Takeshi
;    APPLICANT:  OGURA, Kyozo
;    APPLICANT:  KOYAMA, Tanetoshi
;    TITLE OF INVENTION:  Mutated Farnesyldiphosphate Synthase Capable
;    TITLE OF INVENTION:  Of Synthesizing Geranylgeranyldiphosphate And Gene Coding Ther
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenyon & Kenyon
;      STREET:  1025 Connecticut Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  U.S.
;      ZIP:  20036-5405
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.25" Floppy Disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;      SOFTWARE:  IBM/Word Perfect 6.1 Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/534,910B
;      FILING DATE:  28-SEPT-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-25253
;      FILING DATE:  14-FEB-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Toffenetti, Judith L.
;      REGISTRATION NUMBER:  39,048
;      REFERENCE/DOCKET NUMBER:  77670/398
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)429-1776
;      TELEFAX:  (202)429-0796
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  297 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Bacillus stearothermophilus
US-08-534-910B-9

  Query Match             45.2%;  Score 657;  DB 1;  Length 297;
  Best Local Similarity   47.7%;  Pred. No. 1.8e-56;
  Matches  133;  Conservative   57;  Mismatches   83;  Indels    6;  Gaps    2;

Qy          1 MTNLPMNKLIDEVNNELSVAINKSVM----DTQLEESMLYSLNAGGKRIRPVLLLLTLDS 56
              |  | : : ::|     :  |::: :       :|:::| ||| ||||||||||| |: :
Db          1 MAQLSVEQFLNEQKQAVETALSRYIERLEGPAKLKKAMAYSLEAGGKRIRPLLLLSTVRA 60

Qy         57 LNTEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALL 116
              |  :  :|:  | |:||||||:||||||||:||||| |||| ||||||:||  |||||| ||
Db         61 LGKDPAVGLPVACAIEMIHTHSLIHDDLPSMDNDDLRRGKPTNHKVFGEAMAILAGDGLL 120

Qy        117 TKAFELIS--SDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIH 174
              | ||:||:   |:|:   |:::::::||: |:|  ||| ||  ||: ||: : |  || ||
Db        121 TYAFQLITEIDDERIPPSVRLRLIERLAKAAGPEGMVAGQAADMEGEGKTLTLSELEYIH 180

Qy        175 KTKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKV 234
              : ||| :| ::| : | |   |       |: :: |||: |||:||:||  | | |:||:|
Db        181 RHKTGKMLQYSVHAGALIGGADARQTRELDEFAAHLGLAFQIRDDILDIEGAEEKIGKRV 240

Qy        235 GSDLENNKSTYVSLLGKDGAEDKLTYHRDAAVDELTQID 273
              |||  |||:|| :||   ||::|||:| :||   |   |
Db        241 GSDQSNNKATYPALLSLAGAKEKLTFHIEAAQRHLRNAD 279


RESULT 4
US-08-886-466-2
; Sequence 2, Application US/08886466C
; Patent No. 6040165
; GENERAL INFORMATION:
;  APPLICANT: Narita, Keishi
;  APPLICANT: Ishida, Chika
;  APPLICANT: Takeuchi, Yoshie
;  APPLICANT: Ohto, Chikara
;  APPLICANT: Ohnuma, Shinichi
;  APPLICANT: Nishino, Tokuzo
;  TITLE OF INVENTION: MUTANT PRENYL DIPHOSPHATE SYNTHASE
;  FILE REFERENCE: 77670/494
;  CURRENT APPLICATION NUMBER: US/08/886,466C
;  CURRENT FILING DATE: 1997-07-10
;  EARLIER APPLICATION NUMBER: JP 8-191635
;  EARLIER FILING DATE: 1996-07-03
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 297
;   TYPE: PRT
;   ORGANISM: Bacillus stearothermophilus
US-08-886-466-2

  Query Match             45.0%;  Score 654;  DB 3;  Length 297;
  Best Local Similarity   47.7%;  Pred. No. 3.6e-56;
  Matches  133;  Conservative   55;  Mismatches   85;  Indels    6;  Gaps    2;
```

```
Qy          1 MTNLPMNKLIDEVNNELSVAINKSVM----DTQLEESMLYSLNAGGKRIRPVLLLLTLDS 56
              |  | : : ::|    :  |::: :       :|:::| ||| ||||||||:||| |: :
Db          1 MAQLSVEQFLNEQKQAVETALSRYIERLEGPAKLKKAMAYSLEAGGKRIRPLLLLSTVRA 60

Qy         57 LNTEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALL 116
              |  :  :|:  | |:|||||||||||||:|||| |||| |||||:||  |||||| ||
Db         61 LGKDPAVGLPVACAIEMIHTYSLIHDDLPSMDNDDLRRGKPTNHKVFGEAMAILAGDGLL 120

Qy        117 TKAFELIS--SDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIH 174
              | ||:||:   |:|:   |:::::::||: |:|  ||| ||  ||: ||: : |  || ||
Db        121 TYAFQLITEIDDERIPPSVRLRLIERLAKAAGPEGMVAGQAADMEGEGKTLTLSELEYIH 180

Qy        175 KTKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKV 234
              : ||| :| ::| : | |    |       |: :: |||: |||:||:||  | | |:|| |
Db        181 RHKTGKMLQYSVHAGALIGGADARQTRELDEFAAHLGLAFQIRDDILDIEGAEEKIGKPV 240

Qy        235 GSDLENNKSTYVSLLGKDGAEDKLTYHRDAAVDELTQID 273
              |||  |||:|| :||   ||::|| :| :||   |   |
Db        241 GSDQSNNKATYPALLSLAGAKEKLAFHIEAAQRHLRNAD 279


RESULT 5
US-09-475-304-2
; Sequence 2, Application US/09475304
; Patent No. 6225096
; GENERAL INFORMATION:
;  APPLICANT: Narita, Keishi
;  APPLICANT: Ishida, Chika
;  APPLICANT: Takeuchi, Yoshie
;  APPLICANT: Ohto, Chikara
;  APPLICANT: Ohnuma, Shinichi
;  APPLICANT: Nishino, Tokuzo
;  TITLE OF INVENTION: MUTANT PRENYL DIPHOSPHATE SYNTHASE
;  FILE REFERENCE: 77670/494
;  CURRENT APPLICATION NUMBER: US/09/475,304
;  CURRENT FILING DATE: 1999-12-30
;  EARLIER APPLICATION NUMBER: JP 8-191635
;  EARLIER FILING DATE: 1996-07-03
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 297
;   TYPE: PRT
;   ORGANISM: Bacillus stearothermophilus
US-09-475-304-2

  Query Match             45.0%;  Score 654;  DB 4;  Length 297;
  Best Local Similarity   47.7%;  Pred. No. 3.6e-56;
  Matches  133;  Conservative   55;  Mismatches   85;  Indels    6;  Gaps    2;

Qy          1 MTNLPMNKLIDEVNNELSVAINKSVM----DTQLEESMLYSLNAGGKRIRPVLLLLTLDS 56
              |  | : : ::|    :  |::: :       :|:::| ||| ||||||||:||| |: :
Db          1 MAQLSVEQFLNEQKQAVETALSRYIERLEGPAKLKKAMAYSLEAGGKRIRPLLLLSTVRA 60

Qy         57 LNTEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALL 116
              |  :  :|:  | |:|||||||||||||:|||| |||| |||||:||  |||||| ||
Db         61 LGKDPAVGLPVACAIEMIHTYSLIHDDLPSMDNDDLRRGKPTNHKVFGEAMAILAGDGLL 120

Qy        117 TKAFELIS--SDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIH 174
              | ||:||:   |:|:   |:::::::||: |:|  ||| ||  ||: ||: : |  || ||
Db        121 TYAFQLITEIDDERIPPSVRLRLIERLAKAAGPEGMVAGQAADMEGEGKTLTLSELEYIH 180

Qy        175 KTKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKV 234
              : ||| :| ::| : | |    |       |: :: |||: |||:||:||  | | |:|| |
Db        181 RHKTGKMLQYSVHAGALIGGADARQTRELDEFAAHLGLAFQIRDDILDIEGAEEKIGKPV 240

Qy        235 GSDLENNKSTYVSLLGKDGAEDKLTYHRDAAVDELTQID 273
              |||  |||:|| :||   ||::|| :| :||   |   |
Db        241 GSDQSNNKATYPALLSLAGAKEKLAFHIEAAQRHLRNAD 279


RESULT 6
US-09-101-126-3
; Sequence 3, Application US/09101126
; Patent No. 6316216
; GENERAL INFORMATION:
;  APPLICANT: OHTO, CHIKARA
;  APPLICANT: NAKANE, HIROYUKI
;  APPLICANT: NISHINO, TOKUZO
;  APPLICANT: OHNUMA, SHINICHI
;  APPLICANT: HIROOKA, KAZUTAKE
;  TITLE OF INVENTION: MUTATED PRENYL DIPHOSPHATE SYNTHASES
;  FILE REFERENCE: 77670/566
;  CURRENT APPLICATION NUMBER: US/09/101,126
;  CURRENT FILING DATE: 1999-04-27
;  EARLIER APPLICATION NUMBER: PCT/JP97/03921
;  EARLIER FILING DATE: 1997-10-29
;  EARLIER APPLICATION NUMBER: JP 8-307506
;  EARLIER FILING DATE: 1996-11-05
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 297
;   TYPE: PRT
;   ORGANISM: Bacillus stearothermophilus
;   FEATURE:
;   OTHER INFORMATION: 86-92 is an Asp-rich domain
US-09-101-126-3

  Query Match             45.0%;  Score 654;  DB 4;  Length 297;
  Best Local Similarity   47.7%;  Pred. No. 3.6e-56;
  Matches  133;  Conservative   55;  Mismatches   85;  Indels    6;  Gaps    2;

Qy          1 MTNLPMNKLIDEVNNELSVAINKSVM----DTQLEESMLYSLNAGGKRIRPVLLLLTLDS 56
              |  | : : ::|    :  |::: :       :|:::| ||| ||||||||:||| |: :
Db          1 MAQLSVEQFLNEQKQAVETALSRYIERLEGPAKLKKAMAYSLEAGGKRIRPLLLLSTVRA 60

Qy         57 LNTEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALL 116
              |  :  :|:  | |:|||||||||||||:|||| |||| |||||:||  |||||| ||
Db         61 LGKDPAVGLPVACAIEMIHTYSLIHDDLPSMDNDDLRRGKPTNHKVFGEAMAILAGDGLL 120

Qy        117 TKAFELIS--SDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIH 174
              | ||:||:   |:|:   |:::::::||: |:|  ||| ||  ||: ||: : |  || ||
Db        121 TYAFQLITEIDDERIPPSVRLRLIERLAKAAGPEGMVAGQAADMEGEGKTLTLSELEYIH 180

Qy        175 KTKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKV 234
              : ||| :| ::| : | |    |       |: :: |||: |||:||:||  | | |:|| |
Db        181 RHKTGKMLQYSVHAGALIGGADARQTRELDEFAAHLGLAFQIRDDILDIEGAEEKIGKPV 240

Qy        235 GSDLENNKSTYVSLLGKDGAEDKLTYHRDAAVDELTQID 273
              |||  |||:|| :||   ||::|| :| :||   |   |
Db        241 GSDQSNNKATYPALLSLAGAKEKLAFHIEAAQRHLRNAD 279


RESULT 7
US-08-534-910B-7
; Sequence 7, Application US/08534910B
; Patent No. 5766911
;  GENERAL INFORMATION:
;    APPLICANT:  KOIKE, Ayumi
;    APPLICANT:  OBATA, Shusei
;    APPLICANT:  NISHINO, Tokuzo
;    APPLICANT:  OHNUMA, Shinichi
;    APPLICANT:  NAKAZAWA, Takeshi
;    APPLICANT:  OGURA, Kyozo
;    APPLICANT:  KOYAMA, Tanetoshi
;    TITLE OF INVENTION:  Mutated Farnesyldiphosphate Synthase Capable
;    TITLE OF INVENTION:  Of Synthesizing Geranylgeranyldiphosphate And Gene Coding T
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenyon & Kenyon
;      STREET:  1025 Connecticut Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  U.S.
;      ZIP:  20036-5405
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.25" Floppy Disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;      SOFTWARE:  IBM/Word Perfect 6.1 Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/534,910B
;      FILING DATE:  28-SEPT-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-25253
;      FILING DATE:  14-FEB-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Toffenetti, Judith L.
;      REGISTRATION NUMBER:  39,048
;      REFERENCE/DOCKET NUMBER:  77670/398
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)429-1776
;      TELEFAX:  (202)429-0796
   INFORMATION FOR SEQ ID NO:  7:
     SEQUENCE CHARACTERISTICS:
       LENGTH:  297 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Bacillus stearothermophilus
US-08-534-910B-7

  Query Match             44.9%;  Score 653;  DB 1;  Length 297;
  Best Local Similarity   47.3%;  Pred. No. 4.5e-56;
  Matches  132;  Conservative   56;  Mismatches   85;  Indels    6;  Gaps    2;

Qy          1 MTNLPMNKLIDEVNNELSVAINKSVM----DTQLEESMLYSLNAGGKRIRPVLLLLTLDS 56
              |  | : : ::|    :  |::: :       :::::| ||| |||||||||:||| |: :
Db          1 MAQLSVEQFLNEQKQAVETALSRYIERLEGPAKVKKAMAYSLEAGGKRIRPLLLLSTVQA 60

Qy         57 LNTEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALL 116
              |  :  :|:  | |:|||||||||||||:|||||  |||| |||||:||  |||||| ||
Db         61 LGKDPAVGLPVACAIEMIHTYSLIHDDLPSMDNDDLRRGKPTNHKVFGEAMAILAGDGLL 120

Qy        117 TKAFELIS--SDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIH 174
              | ||:||:   |:|:   |:::::::||: |:|  ||| ||  ||: ||: : |  || ||
Db        121 TYAFQLITEIDDERIPPSVRLRLIERLAKAAGPEGMVAGQAADMEGEGKTLTLSELEYIH 180

Qy        175 KTKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKV 234
              : ||| :| ::| : | |   |      |: :: ||| ||||:||:||  | | |:|| |
Db        181 RHKTGKMLQYSVHAGALIGGADARQTRELDEFAAHLGLAFQIRDDILDIEGAEEKIGKPV 240

Qy        235 GSDLENNKSTYVSLLGKDGAEDKLTYHRDAAVDELTQID 273
              |||  |||:|| :||   ||::|| :| :||   |   |
Db        241 GSDQSNNKATYPALLSLAGAKEKLAFHIEAAQRHLRNAD 279


RESULT 8
US-08-534-910B-8
; Sequence 8, Application US/08534910B
; Patent No. 5766911
;  GENERAL INFORMATION:
;    APPLICANT:  KOIKE, Ayumi
;    APPLICANT:  OBATA, Shusei
;    APPLICANT:  NISHINO, Tokuzo
;    APPLICANT:  OHNUMA, Shinichi
;    APPLICANT:  NAKAZAWA, Takeshi
;    APPLICANT:  OGURA, Kyozo
;    APPLICANT:  KOYAMA, Tanetoshi
;    TITLE OF INVENTION:  Mutated Farnesyldiphosphate Synthase Capable
;    TITLE OF INVENTION:  Of Synthesizing Geranylgeranyldiphosphate And Gene Coding Ther
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenyon & Kenyon
;      STREET:  1025 Connecticut Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  U.S.
;      ZIP:  20036-5405
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.25" Floppy Disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;      SOFTWARE:  IBM/Word Perfect 6.1 Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/534,910B
;      FILING DATE:  28-SEPT-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-25253
;      FILING DATE:  14-FEB-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Toffenetti, Judith L.
;      REGISTRATION NUMBER:  39,048
;      REFERENCE/DOCKET NUMBER:  77670/398
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)429-1776
;      TELEFAX:  (202)429-0796
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  297 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Bacillus stearothermophilus
US-08-534-910B-8

  Query Match             44.7%;  Score 650;  DB 1;  Length 297;
  Best Local Similarity   47.3%;  Pred. No. 8.9e-56;
  Matches  132;  Conservative   55;  Mismatches   86;  Indels    6;  Gaps    2;

Qy          1 MTNLPMNKLIDEVNNELSVAINKSVM----DTQLEESMLYSLNAGGKRIRPVLLLLTLDS 56
              |  | : : ::|    :  |::: :       :|:::| ||| |||||||||:||| |: :
Db          1 MAQLSVEQFLNEQKQAVETALSRYIERLEGPAKLKKAMAYSLEAGGKRIRPLLLLSTVRA 60

Qy         57 LNTEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALL 116
              |  :  :|:  | |:|||||||||||||:|||||  |||| |||||:||  |||||| ||
Db         61 LGKDPAVGLPVACAIEMIHTYSLIHDDLPSMDNDDLRRGKPTNHKVFGEAMAILAGDGLL 120

Qy        117 TKAFELIS--SDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIH 174
              | ||:||:   |:|:   |:::::::||: |:|  ||  ||  ||: ||: : |  || ||
Db        121 TYAFQLITEIDDERIPPSVRLRLIERLAKAAGPEGMAAGQAADMEGEGKTLTLSELEYIH 180

Qy        175 KTKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKV 234
              : ||| :| ::| : | |   |      |: :: ||| |||:||:||  | | |:|| |
Db        181 RYKTGKMLQYSVHAGALIGGADARQTRELDEFAAHLGLAFQIRDDILDIEGAEEKIGKPV 240

Qy        235 GSDLENNKSTYVSLLGKDGAEDKLTYHRDAAVDELTQID 273
              |||  |||:|| :||   ||::|| :| :||   |   |
Db        241 GSDQSNNKATYPALLSLAGAKEKLAFHIEAAQRHLRNAD 279


RESULT 9
US-08-534-910B-10
; Sequence 10, Application US/08534910B
; Patent No. 5766911
;  GENERAL INFORMATION:
;    APPLICANT:  KOIKE, Ayumi
;    APPLICANT:  OBATA, Shusei
;    APPLICANT:  NISHINO, Tokuzo
;    APPLICANT:  OHNUMA, Shinichi
;    APPLICANT:  NAKAZAWA, Takeshi
;    APPLICANT:  OGURA, Kyozo
;    APPLICANT:  KOYAMA, Tanetoshi
;    TITLE OF INVENTION:  Mutated Farnesyldiphosphate Synthase Capable
;    TITLE OF INVENTION:  Of Synthesizing Geranylgeranyldiphosphate And Gene Coding T
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
```

```
;      ADDRESSEE:  Kenyon & Kenyon
;      STREET:  1025 Connecticut Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  U.S.
;      ZIP:  20036-5405
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.25" Floppy Disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;      SOFTWARE:  IBM/Word Perfect 6.1 Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/534,910B
;      FILING DATE:  28-SEPT-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-25253
;      FILING DATE:  14-FEB-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Toffenetti, Judith L.
;      REGISTRATION NUMBER:  39,048
;      REFERENCE/DOCKET NUMBER:  77670/398
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)429-1776
;      TELEFAX:  (202)429-0796
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  297 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Bacillus stearothermophilus
US-08-534-910B-10

  Query Match             44.7%;  Score 650;  DB 1;  Length 297;
  Best Local Similarity   47.3%;  Pred. No. 8.9e-56;
  Matches  132;  Conservative   55;  Mismatches   86;  Indels    6;  Gaps    2;

Qy          1 MTNLPMNKLIDEVNNELSVAINKSVM----DTQLEESMLYSLNAGGKRIRPVLLLLTLDS 56
              |  | : : ::|    :  |::: :        : :::| ||| ||||||||:||| |: :
Db          1 MAQLSVEQFLNEQKQAVETALSRYIERLEGPAKKKKAMAYSLEAGGKRIRPLLLLSTVQA 60

Qy         57 LNTEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALL 116
              |  : .:|:  | |:||||||||||||||:|||| |||| |||||:||  |||||| ||
Db         61 LGKDPAVGLPVACAIEMIHTYSLIHDDLPSMDNDDLRRGKPTNHKVFGEAMAILAGDGLL 120

Qy        117 TKAFELIS--SDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIH 174
              | ||:||:   |:|:   |:::::::||: |:|  ||| ||  ||: ||: : |  || ||
Db        121 TYAFQLITEIDDERIPPSVRLRLIERLAKAAGPEGMVAGQAADMEGEGKTLTLSELEYIH 180

Qy        175 KTKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKV 234
              : ||| :| ::| : | |   |        |: :: ||| ||| ||:||:||  | | |:|| |
Db        181 RHKTGKMLQYSVHAGALIGGADARQTRELDEFAAHLGLAFQIRDDILDIEGAEEKIGKPV 240

Qy        235 GSDLENNKSTYVSLLGKDGAEDKLTYHRDAAVDELTQID 273
              |||  |||:|| :||   ||::|| :| :||   |   |
Db        241 GSDQSNNKATYPALLSLAGAKEKLAFHIEAAQRHLRNAD 279


RESULT 10
US-09-367-528A-5
; Sequence 5, Application US/09367528A
; Patent No. 6395525
; GENERAL INFORMATION:
;  APPLICANT:  TOYOTA JIDOSHA KABUSHIKI KAISHA
;  TITLE OF INVENTION:  Geranyl Diphosphate Synthetase Gene
;  FILE REFERENCE:  PH-586
;  CURRENT APPLICATION NUMBER: US/09/367,528A
;  CURRENT FILING DATE:  1999-08-16
;  PRIOR APPLICATION NUMBER:  JP97/346686
;  PRIOR FILING DATE:  1997-12-16
;  NUMBER OF SEQ ID NOS:  6
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 297
;   TYPE: PRT
;   ORGANISM: Bacillus stearothermophilus
US-09-367-528A-5

  Query Match             44.7%;  Score 650;  DB 4;  Length 297;
  Best Local Similarity   49.6%;  Pred. No. 8.9e-56;
  Matches  134;  Conservative   49;  Mismatches   85;  Indels    2;  Gaps    1;

Qy          6 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGM 65
              :|:    |   ||  | :    :|:::| ||| ||||||||:||| |: :|  :  :|:
Db         10 LNEQKQAVETALSRYIERLEGPAKLKKAMAYSLEAGGKRIRPLLLLSTVRALGKDPAVGL 69

Qy         66 KSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELIS- 124
               | |:||||||||||||||:|||| |||| |||||:||  |||||| ||| ||:||:
Db         70 PVACAIEMIHTYSLIHDDLPSMDNDDLRRGKPTNHKVFGEAMAILAGDGLLTYAFQLITE 129

Qy        125 -SDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLT 183
               |:|:   |:::::::||: |:|  ||| ||  ||: ||: : |  || ||: ||| :|
Db        130 IDDERIPPSVRLRLIERLAKAAGPEGMVAGQAADMEGEGKTLTLSELEYIHRHKTGKMLQ 189

Qy        184 FAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKS 243
              ::| : | |   |       |: :: ||| |||:||:||  | | |:|| |||  |||:
Db        190 YSVHAGALIGGADARQTRELDEFAAHLGLAFQIRDDILDIEGAEEKIGKPVGSDQSNNKA 249

Qy        244 TYVSLLGKDGAEDKLTYHRDAAVDELTQID 273
              || :||   ||::|| :| :||   |   |
Db        250 TYPALLSLAGAKEKLAFHIEAAQRHLRNAD 279


RESULT 11
US-09-367-528A-1
; Sequence 1, Application US/09367528A
; Patent No. 6395525
; GENERAL INFORMATION:
;  APPLICANT:  TOYOTA JIDOSHA KABUSHIKI KAISHA
;  TITLE OF INVENTION:  Geranyl Diphosphate Synthetase Gene
;  FILE REFERENCE:  PH-586
;  CURRENT APPLICATION NUMBER: US/09/367,528A
;  CURRENT FILING DATE:  1999-08-16
;  PRIOR APPLICATION NUMBER:  JP97/346686
;  PRIOR FILING DATE:  1997-12-16
;  NUMBER OF SEQ ID NOS:  6
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 297
;   TYPE: PRT
;   ORGANISM: Bacillus stearothermophilus
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: 82
;   OTHER INFORMATION: Xaa represents Val, Leu, Ile, Thr, Asp, Glu, Asn, Gln, Lys,
;   OTHER INFORMATION: Arg, Cys, Met, Phe, Tyr, Trp, His or Pro.
US-09-367-528A-1

  Query Match             44.5%;  Score 646;  DB 4;  Length 297;
  Best Local Similarity   49.3%;  Pred. No. 2.2e-55;
  Matches  133;  Conservative   49;  Mismatches   86;  Indels    2;  Gaps    1;

Qy          6 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGM 65
              :|:    |   ||  | :    :|:::| ||| ||||||||:||| |: :|  :  :|:
Db         10 LNEQKQAVETALSRYIERLEGPAKLKKAMAYSLEAGGKRIRPLLLLSTVRALGKDPAVGL 69

Qy         66 KSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELIS- 124
               | |:|||||| |||||||:|||| |||| |||||:||  |||||| ||| ||:||:
Db         70 PVACAIEMIHTYXLIHDDLPSMDNDDLRRGKPTNHKVFGEAMAILAGDGLLTYAFQLITE 129

Qy        125 -SDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLT 183
               |:|:   |:::::::||: |:|  ||| ||  ||: ||: : |  || ||: ||| :|
Db        130 IDDERIPPSVRLRLIERLAKAAGPEGMVAGQAADMEGEGKTLTLSELEYIHRHKTGKMLQ 189
```

```
Qy        184 FAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKS 243
              ::| : | |    |      |: :: |||: |||:||:||  | | |:|| ||||  |||:
Db        190 YSVHAGALIGGADARQTRELDEFAAHLGLAFQIRDDILDIEGAEEKIGKPVGSDQSNNKA 249

Qy        244 TYVSLLGKDGAEDKLTYHRDAAVDELTQID 273
              || :||   ||::|| :| :||   |   |
Db        250 TYPALLSLAGAKEKLAFHIEAAQRHLRNAD 279


RESULT 12
US-08-534-910B-6
; Sequence 6, Application US/08534910B
; Patent No. 5766911
;  GENERAL INFORMATION:
;    APPLICANT:  KOIKE, Ayumi
;    APPLICANT:  OBATA, Shusei
;    APPLICANT:  NISHINO, Tokuzo
;    APPLICANT:  OHNUMA, Shinichi
;    APPLICANT:  NAKAZAWA, Takeshi
     APPLICANT:  OGURA, Kyozo
     APPLICANT:  KOYAMA, Tanetoshi
     TITLE OF INVENTION:  Mutated Farnesyldiphosphate Synthase Capable
;    TITLE OF INVENTION:  Of Synthesizing Geranylgeranyldiphosphate And Gene Coding Ther
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenyon & Kenyon
;      STREET:  1025 Connecticut Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  U.S.
;      ZIP:  20036-5405
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.25" Floppy Disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;      SOFTWARE:  IBM/Word Perfect 6.1 Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/534,910B
;      FILING DATE:  28-SEPT-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-25253
;      FILING DATE:  14-FEB-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Toffenetti, Judith L.
;      REGISTRATION NUMBER:  39,048
;      REFERENCE/DOCKET NUMBER:  77670/398
     TELECOMMUNICATION INFORMATION:
       TELEPHONE:  (202)429-1776
       TELEFAX:  (202)429-0796
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  297 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Bacillus stearothermophilus
US-08-534-910B-6

  Query Match             44.4%;  Score 645;  DB 1;  Length 297;
  Best Local Similarity   48.0%;  Pred. No. 2.8e-55;
  Matches  130;  Conservative   56;  Mismatches   79;  Indels    6;  Gaps    2;

Qy          1 MTNLPMNKLIDEVNNELSVAINKSVM----DTQLEESMLYSLNAGGKRIRPVLLLLTLDS 56
              |  | : : ::|    :  |::: :       :|:::| ||| ||||||||:||| |: :
Db          1 MAQLSVEQFLNEQKQAVETALSRYIERLEGPAKLKKAMAYSLEAGGKRIRPLLLLSTVRA 60

Qy         57 LNTEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALL 116
              |  :  :|:  | |:|||||:||||||||:|||||  |||| |||||:||  |||||| ||
Db         61 LGKDPAVGLPVACAIEMIHTHSLIHDDLPSMDNDDLRRGKPTNHKVFGEAMAILAGDGLL 120

Qy        117 TKAFELIS--SDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIH 174
              | ||:||:   |:|:   |:::::::||: |:|  ||| ||  ||: ||: : |  || ||
Db        121 TYAFQLITEIDDERIPPSVRLRLIERLAKAAGPEGMVAGQAADMEGEGKTLTLSELEYIH 180

Qy        175 KTKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKV 234
              : ||| :| ::| : | |    |       |: :: |||: |||:||:||  | | |:|| |
Db        181 RHKTGKMLQYSVHAGALIGGADARQTRELDEFAAHLGLAFQIRDDILDIEGAEEKIGKPV 240

Qy        235 GSDLENNKSTYVSLLGKDGAEDKLTYHRDAA 265
              |||  |||:|| :||   ||::|| :| :||
Db        241 GSDQSNNKATYPALLSLAGAKEKLAFHIEAA 271


RESULT 13
US-09-367-528A-3
; Sequence 3, Application US/09367528A
; Patent No. 6395525
; GENERAL INFORMATION:
;  APPLICANT:  TOYOTA JIDOSHA KABUSHIKI KAISHA
;  TITLE OF INVENTION:  Geranyl Diphosphate Synthetase Gene
;  FILE REFERENCE:  PH-586
;  CURRENT APPLICATION NUMBER: US/09/367,528A
;  CURRENT FILING DATE:  1999-08-16
;  PRIOR APPLICATION NUMBER:  JP97/346686
;  PRIOR FILING DATE:  1997-12-16
;  NUMBER OF SEQ ID NOS:  6
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 297
;   TYPE: PRT
;   ORGANISM: Bacillus stearothermophilus
US-09-367-528A-3

  Query Match             44.3%;  Score 644;  DB 4;  Length 297;
  Best Local Similarity   49.3%;  Pred. No. 3.5e-55;
  Matches  133;  Conservative   49;  Mismatches   86;  Indels    2;  Gaps    1;

Qy          6 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGM 65
              :|:    |   ||  | :     :|:::| ||| |||||||||:||| |: :|  :  :|:
Db         10 LNEQKQAVETALSRYIERLEGPAKLKKAMAYSLEAGGKRIRPLLLLSTVRALGKDPAVGL 69

Qy         66 KSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELIS- 124
               | |:|||||| |||||||:|||| |||| |||||:||  |||||| ||| ||:||:
Db         70 PVACAIEMIHTYFLIHDDLPSMDNDDLRRGKPTNHKVFGEAMAILAGDGLLTYAFQLITE 129

Qy        125 -SDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLT 183
               |:|:   |::::::||: |:|  ||| ||  ||: ||: : |  || ||: ||| :|
Db        130 IDDERIPPSVRLRLIERLAKAAGPEGMVAGQAADMEGEGKTLTLSELEYIHRHKTGKMLQ 189

Qy        184 FAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKS 243
              ::| : | |    |      |: :: |||: |||:||:||  | | |:|| ||||  |||:
Db        190 YSVHAGALIGGADARQTRELDEFAAHLGLAFQIRDDILDIEGAEEKIGKPVGSDQSNNKA 249

Qy        244 TYVSLLGKDGAEDKLTYHRDAAVDELTQID 273
              || :||   ||::|| :| :||   |   |
Db        250 TYPALLSLAGAKEKLAFHIEAAQRHLRNAD 279


RESULT 14
US-09-275-742-2
; Sequence 2, Application US/09275742
; Patent No. 6130069
; GENERAL INFORMATION:
;  APPLICANT: Wilding, Edwina Imogen
;  APPLICANT: Gwynn, Michael
;  TITLE OF INVENTION: ispA
;  FILE REFERENCE: GM10205
;  CURRENT APPLICATION NUMBER: US/09/275,742
;  CURRENT FILING DATE: 1999-03-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 291
;   TYPE: PRT
```

```
;  ORGANISM: Streptococcus pneumoniae
US-09-275-742-2

  Query Match             37.9%;  Score 551;  DB 4;  Length 291;
  Best Local Similarity   46.6%;  Pred. No. 4.2e-46;
  Matches  122;  Conservative   48;  Mismatches   88;  Indels    4;  Gaps    3;

Qy         28 TQLEESMLYSLNAGGKRIRPVLLLLTLDSLN-TEYELGMKSAIALEMIHTYSLIHDDLPA 86
              : | ||:|||::|||||||| ||| |::|  |      : | ||||||| |||||||||
Db         25 SSLRESVLYSIHAGGKRIRPFLLLKVLEALQVTIKPAHAQVATALEMIHTGSLIHDDLPA 84

Qy         87 MDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSDDRLTDEVKIKVLQRLSIASG 146
              ||:||||||:||||| :||  |||||||    : ||:  | |  ::|: ::  ||:|||
Db         85 MDDDDYRRGRLTNHKKFGEAMAILAGDALFLDPYALIAQAD-LPSQIKVGLIANLSLASG 143

Qy        147 HVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFAVMSAADIANVDDTTKEHLESY 206
               :||| ||:|||: | | : || |: ||  ||| || :   :|| || :    :  |::
Db        144 SLGMVAGQVLDMEGEHQHLSLEELQTIHANKTGKLLAYPFQAAAIIAELSPEMQVKLKTV 203

Qy        207 SYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVSLLGKDGAEDKLTYHRDAAV 266
              :|: ||::||:||      ::||    ||:  |||| :||| : :        | |
Db        204 GELIGLAFQVRDDVLDVTASFEEIGKTPQKDLQAEKSTYPALLGLEESIAFCNQTLDQAN 263

Qy        267 DELTQIDEQ--FNTKHLLEIVD 286
              |:| :| :|  | |: :: :|:
Db        264 DKLEEIAQQLPFETESIVSVVE 285


RESULT 15
US-09-187-050-34
; Sequence 34, Application US/09187050B
; Patent No. 6043072
; GENERAL INFORMATION:
;  APPLICANT: Croteau, Rodney B
;  APPLICANT: Hefner, Jerry
;  TITLE OF INVENTION: Nucleic Acids Encoding Taxus Geranylgeranyl Diphosphate
;  TITLE OF INVENTION: Synthase, And Methods of Use
;  FILE REFERENCE: WSUR12423
;  CURRENT APPLICATION NUMBER: US/09/187,050B
;  CURRENT FILING DATE: 1998-11-05
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
;   LENGTH: 393
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Taxus GGPP
;   OTHER INFORMATION: synthase protein variant
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (1)..(393)
;   OTHER INFORMATION: Computer-generated Taxus GGPP synthase protein
;   OTHER INFORMATION: variant
US-09-187-050-34

  Query Match             35.9%;  Score 521;  DB 3;  Length 393;
  Best Local Similarity   38.3%;  Pred. No. 5.7e-43;
  Matches  118;  Conservative   55;  Mismatches  107;  Indels   28;  Gaps    5;

Qy          6 MNKLIDEVNNE--------------------LSVAINKSV---MDTQLEESMLYSLNAG 41
              | :|:|    |                    :  |::|::      :: ||| ||| ||
Db         80 MGQLVDLAETEKAEGKDIEFDFNEYMKSKAVAVDAALDKAIPLEYPEKIHESMRYSLLAG 139

Qy         42 GKRIRPVLLLLTLDSLNTEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHK 101
              |||:|| | :   : :    :| | :| |:|||| |||||||| |||||:|||| ||||
Db        140 GKRVRPALCIAACELVGGSQDLAMPTACAMEMIHTMSLIHDDLPCMDNDDFRRGKPTNHK 199

Qy        102 VYGEWTAILAGDALLTKAFE--LISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQ 159
              |:|| ||:|||||||: |||   :::   :  :  ::|:  |    |  |:||||::|:
Db        200 VFGEDTAVLAGDALLSFAFEHIAVATSKTVPSDRTLRVISELGKTIGSQGLVGGQVVDIT 259

Qy        160 SEGQP-IDLETLEMIHKTKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKD 218
                 |||   :||:||| ||  ||  ||  :|:|   :    :      :  |:  :|::||: |
Db        260 SEGDANVDLKTLEWIHIHKTAVLLECSVVSGGILGGATEDEIARIRRYARCVGLLFQVVD 319

Qy        219 DLLDCYGDEAKLGKKVGSDLENNKSTYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNT 278
              |:||      :|||  | ||  :|:|| |:| : |::       | :||:  | |
Db        320 DILDVTKSSEELGKTAGKDLLTDKATYPKLMGLEKAKEFAAELATRAKEELSSFD-QIKA 378

Qy        279 KHLLEIVD 286
                || : |
Db        379 APLLGLAD 386


Search completed: May 29, 2003, 12:00:37
Job time : 17 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 29, 2003, 11:59:55 ; Search time 45 Seconds
(without alignments)
645.582 Million cell updates/sec

Title: US-09-925-637-64
Perfect score: 1453
Sequence: 1 MTNLPMNKLIDEVNNELSVA.....ELTQIDEQFNTKHLLEIVDL 287

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 383519 seqs, 101223694 residues

Total number of hits satisfying chosen parameters: 383519

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1453	100.0	287	9	US-10-084-205-64	Sequence 64, Appl
2	1453	100.0	287	10	US-09-925-637-64	Sequence 64, Appl
3	1442	99.2	293	10	US-09-815-242-12583	Sequence 12583, A
4	1415	97.4	288	10	US-09-815-242-5239	Sequence 5239, Ap
5	596	41.0	293	10	US-09-815-242-10630	Sequence 10630, A
6	549	37.8	291	10	US-09-815-242-13597	Sequence 13597, A
7	547	37.6	291	10	US-09-815-242-13273	Sequence 13273, A
8	516.5	35.5	297	9	US-09-941-947A-20	Sequence 20, Appl
9	516.5	35.5	297	10	US-09-934-903-14	Sequence 14, Appl
10	516.5	35.5	297	10	US-09-934-868-72	Sequence 72, Appl
11	504.5	34.7	367	9	US-10-108-915-26	Sequence 26, Appl
12	497.5	34.2	295	10	US-09-815-242-11239	Sequence 11239, A
13	494.5	34.0	350	9	US-10-108-915-22	Sequence 22, Appl
14	487.5	33.6	369	9	US-10-108-915-18	Sequence 18, Appl
15	485.5	33.4	299	10	US-09-815-242-10069	Sequence 10069, A
16	478.5	32.9	316	9	US-10-108-915-45	Sequence 45, Appl
17	474.5	32.7	295	10	US-09-815-242-11971	Sequence 11971, A
18	473	32.6	299	10	US-09-815-242-14084	Sequence 14084, A
19	465.5	32.0	377	10	US-09-934-778-2	Sequence 2, Appli
20	402.5	27.7	316	9	US-10-108-915-16	Sequence 16, Appl
21	402.5	27.7	316	9	US-10-108-915-20	Sequence 20, Appl
22	376.5	25.9	303	9	US-09-941-947A-26	Sequence 26, Appl
23	373.5	25.7	303	10	US-09-815-242-11382	Sequence 11382, A
24	372.5	25.6	303	10	US-09-815-242-11547	Sequence 11547, A
25	356	24.5	232	9	US-10-108-915-14	Sequence 14, Appl
26	350	24.1	228	9	US-10-108-915-12	Sequence 12, Appl
27	311.5	21.4	322	10	US-09-815-242-12603	Sequence 12603, A
28	311.5	21.4	342	9	US-10-108-915-38	Sequence 38, Appl
29	311	21.4	323	10	US-09-815-242-10321	Sequence 10321, A
30	310	21.3	323	10	US-09-815-242-13734	Sequence 13734, A
31	309.5	21.3	295	9	US-09-920-923-2	Sequence 2, Appli
32	309.5	21.3	295	10	US-09-547-267-1	Sequence 1, Appli
33	308.5	21.2	312	10	US-09-815-242-5315	Sequence 5315, Ap
34	303	20.9	326	9	US-10-108-915-46	Sequence 46, Appl
35	300	20.6	293	9	US-09-920-923-37	Sequence 37, Appl
36	294	20.2	329	10	US-09-815-242-11142	Sequence 11142, A
37	292	20.1	209	9	US-10-108-915-30	Sequence 30, Appl
38	280	19.3	322	10	US-09-815-242-12031	Sequence 12031, A
39	269	18.5	342	9	US-10-108-915-34	Sequence 34, Appl
40	260	17.9	345	9	US-10-108-915-44	Sequence 44, Appl
41	237.5	16.3	273	9	US-10-084-205-66	Sequence 66, Appl
42	237.5	16.3	273	10	US-09-925-637-66	Sequence 66, Appl
43	230.5	15.9	371	9	US-09-738-626-5880	Sequence 5880, Ap
44	223	15.3	326	9	US-09-252-088-35	Sequence 35, Appl
45	223	15.3	350	9	US-09-738-626-4030	Sequence 4030, Ap

ALIGNMENTS

RESULT 1
US-10-084-205-64
; Sequence 64, Application US/10084205
; Publication No. US20030049648A1
; GENERAL INFORMATION:
; APPLICANT: Choi, Gil
; TITLE OF INVENTION: 37 Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB515P1
; CURRENT APPLICATION NUMBER: US/10/084,205
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/151,933
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 64
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-084-205-64

Query Match 100.0%; Score 1453; DB 9; Length 287;
Best Local Similarity 100.0%; Pred. No. 1.3e-119;
Matches 287; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MTNLPMNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTNLPMNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTE 60

Qy         61 YELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAF 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAF 120

Qy        121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGA 180

Qy        181 LLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLEN 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLEN 240
```

```
Qy        241 NKSTYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NKSTYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287


RESULT 2
US-09-925-637-64
; Sequence 64, Application US/09925637
; Patent No. US20020103338A1
; GENERAL INFORMATION:
;  APPLICANT: Choi
;  TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
;  FILE REFERENCE: PB560
;  CURRENT APPLICATION NUMBER: US/09/925,637
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/23773
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US 60/151,933
;  PRIOR FILING DATE: 1999-09-01
;  PRIOR APPLICATION NUMBER: US 08/781,986
;  PRIOR FILING DATE: 1997-01-03
;  PRIOR APPLICATION NUMBER: US 08/956,171
;  PRIOR FILING DATE: 1997-10-20
;  PRIOR APPLICATION NUMBER: US 60/009,861
;  PRIOR FILING DATE: 1996-01-06
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 64
;   LENGTH: 287
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-925-637-64

  Query Match           100.0%;  Score 1453;  DB 10;  Length 287;
  Best Local Similarity  100.0%;  Pred. No. 1.3e-119;
  Matches  287;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MTNLPMNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTNLPMNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTE 60

Qy         61 YELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAF 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAF 120

Qy        121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGA 180

          181 LLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLEN 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLEN 240

Qy        241 NKSTYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NKSTYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287


RESULT 3
US-09-815-242-12583
; Sequence 12583, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12583
;   LENGTH: 293
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-09-815-242-12583

  Query Match           99.2%;  Score 1442;  DB 10;  Length 293;
  Best Local Similarity  99.3%;  Pred. No. 1.2e-118;
  Matches  285;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MTNLPMNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTNLPMNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTE 60

Qy         61 YELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAF 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAF 120

Qy        121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGA 180

Qy        181 LLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLEN 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        181 LLTFAVMSAADIANVDDATKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLEN 240

Qy        241 NKSTYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287
              ||||||||||||||||||||||||||||| ||||||||||||||||||
Db        241 NKSTYVSLLGKDGAEDKLTYHRDAAVDELMQIDEQFNTKHLLEIVDL 287


RESULT 4
US-09-815-242-5239
; Sequence 5239, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
```

```
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5239
;   LENGTH: 288
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-09-815-242-5239

  Query Match           97.4%;  Score 1415;  DB 10;  Length 288;
  Best Local Similarity   99.3%;  Pred. No. 2.8e-116;
  Matches  280;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          6 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGM 65
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGM 60

Qy         66 KSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISS 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISS 120

Qy        126 DDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFA 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFA 180

Qy        186 VMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTY 245
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VMSAADIANVDDATKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTY 240

Qy        246 VSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287
              ||||||||||||||||||||||| ||||||||||||||||
Db        241 VSLLGKDGAEDKLTYHRDAAVDELMQIDEQFNTKHLLEIVDL 282


RESULT 5
US-09-815-242-10630
; Sequence 10630, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
   APPLICANT:  Ohlsen, Kari L.
   APPLICANT:  Zyskind, Judith W.
   APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10630
;   LENGTH: 293
;   TYPE: PRT
;   ORGANISM: Enterococcus faecalis
US-09-815-242-10630

  Query Match           41.0%;  Score 596;  DB 10;  Length 293;
  Best Local Similarity   46.0%;  Pred. No. 1.7e-44;
  Matches  134;  Conservative   56;  Mismatches   91;  Indels   10;  Gaps    5;

Qy          1 MTNLPMN--KLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLN 58
              |||       |:::|  :    | :   : :|:|:||||::|||||:||:|:| |: :
Db          1 MTNFSQQHLPLVEKVMVDF---IAEYTENERLKEAMLYSIHAGGKRLRPLLVLTTVAAFQ 57

Qy         59 TEYEL-GMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLT 117
               | |    : | :|||||||||||||||||:|| |||| |||||:|| |||||| |||
Db         58 KEMETQDYQVAASLEMIHTYSLIHDDLPAMDDDDLRRGKPTNHKVFGEATAILAGDGLLT 117

Qy        118 KAFELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTK 177
              ||:|:|       | |: ::|:|: |:|: ||| ||| |:: |   : || |  :|: |
Db        118 GAFQLLSLSQLGLSE-KVLLMQQLAKAAGNQGMVAGQMGDIEGEKVSLTLEELAAVHEKK 176

Qy        178 TGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSD 237
              ||||: ||:::   :||  :     |  :::| |: |||:|||||    || |||||| |
Db        177 TGALIEFALIAGGVLANQTEEVIGLLTQFAHHYGLAFQIRDDLLDATSTEADLGKKVGRD 236

Qy        238 LENNKSTYVSLLGKDGAEDKLTY---HRDAAVDELTQIDEQFNTKHLLEIV 285
              ||||| :|||  ||:| ||:      | ::::      |: :||  ::
Db        237 EALNKSTYPALLGIAGAKDALTHQLAEGSAVLEKIKANVPNFSEEHLANLL 287


RESULT 6
US-09-815-242-13597
; Sequence 13597, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13597
;   LENGTH: 291
;   TYPE: PRT
;   ORGANISM: Streptococcus pneumoniae
US-09-815-242-13597
```

```
  Query Match           37.8%;  Score 549;  DB 10;  Length 291;
  Best Local Similarity  46.2%;  Pred. No. 2.3e-40;
  Matches  121;  Conservative   50;  Mismatches   87;  Indels    4;  Gaps    3;

Qy         28 TQLEESMLYSLNAGGKRIRPVLLLLTLDSLN-TEYELGMKSAIALEMIHTYSLIHDDLPA 86
              : | ||:|||::|||||||| |||  |::|  |      : | ||||||| ||||||||
Db         25 SSLRESVLYSIHAGGKRIRPFLLLEVLEALQVTIKPAHAQVATALEMIHTGSLIHDDLPA 84

Qy         87 MDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSDDRLTDEVKIKVLQRLSIASG 146
              ||:||||||:||||| :||  ||||||||    : ||:  | |  ::|: ::  ||:|||
Db         85 MDDDDYRRGRLTNHKKFGEAMAILAGDALFLDPYALIAQAD-LPSQIKVDLIANLSLASG 143

Qy        147 HVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFAVMSAADIANVDDTTKEHLESY 206
               :||| ||:|||: | | : || |::||  ||| || :   :|| || :    :  |::
Db        144 SLGMVAGQVLDMEGEHQHLSLEELQIIHANKTGKLLAYPFQAAAIIAELSPEMQVKLKTV 203

Qy        207 SYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVSLLGKDGAEDKLTYHRDAAV 266
               :|: ||::||:||      ::||    ||:  |||| :||| : :        | |
Db        204 GELIGLAFQVRDDVLDVTASFEEIGKTPQKDLQAEKSTYPALLGLEESIAFCNQTLDQAN 263

Qy        267 DELTQIDEQ--FNTKHLLEIVD 286
              ::| :| :|  | |: :: :|:
Db        264 EKLEEIAQQLPFETESIVSVVE 285


RESULT 7
US-09-815-242-13273
; Sequence 13273, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
   PRIOR FILING DATE: 2000-05-23
   PRIOR APPLICATION NUMBER: 60/207,727
   PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13273
;   LENGTH: 291
;   TYPE: PRT
;   ORGANISM: Streptococcus pneumoniae
US-09-815-242-13273

  Query Match           37.6%;  Score 547;  DB 10;  Length 291;
  Best Local Similarity  45.8%;  Pred. No. 3.4e-40;
  Matches  120;  Conservative   50;  Mismatches   88;  Indels    4;  Gaps    3;

Qy         28 TQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYE-LGMKSAIALEMIHTYSLIHDDLPA 86
              : | ||:|||::|||||||| |||  |::|    :    : | ||||||| ||||||||
Db         25 SSLRESVLYSIHAGGKRIRPFLLLEVLEALQVAIKPAHAQVATALEMIHTGSLIHDDLPA 84

Qy         87 MDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSDDRLTDEVKIKVLQRLSIASG 146
              ||:||||||:||||| :||  ||||||||    : ||:  | |  ::|: ::  ||:|||
Db         85 MDDDDYRRGRLTNHKKFGEAMAILAGDALFLDPYALIAQAD-LPSQIKVDLIANLSLASG 143

Qy        147 HVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFAVMSAADIANVDDTTKEHLESY 206
               :||| ||:|||: | | : || |::||  ||| || :   :|| || :    :  |::
Db        144 SLGMVAGQVLDMEGEHQHLSLEELQTIHANKTGKLLAYPFQAAAIIAELSPEMQVKLKTV 203

Qy        207 SYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVSLLGKDGAEDKLTYHRDAAV 266
               :|: ||::||:||      ::||    ||:  |||| :||| : :        | |
Db        204 GELIGLAFQVRDDVLDVTASFEEIGKTPQKDLQAEKSTYPALLGLEESIAFCNQTLDQAN 263

Qy        267 DELTQIDEQ--FNTKHLLEIVD 286
              ::| :| :|  | |: :: :|:
Db        264 EKLEEIAQQVPFETESIVSVVE 285


RESULT 8
US-09-941-947A-20
; Sequence 20, Application US/09941947A
; Publication No. US20030003528A1
; GENERAL INFORMATION:
;  APPLICANT: Brzostowicz, Patricia C.
;  APPLICANT:  Cheng, Qiong
;  APPLICANT:  DiCosimo, Deana J.
;  APPLICANT:  Koffas, Mattheos
;  APPLICANT:  Miller, Edward S. Jr.
;  APPLICANT:  Odom, J. Martin
;  APPLICANT:  Picataggio, Steve
;  APPLICANT:  Rouviere, Pierre E.
;  TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
;  FILE REFERENCE: CL1903 US NA
;  CURRENT APPLICATION NUMBER: US/09/941,947A
;  CURRENT FILING DATE:  2001-09-01
;  PRIOR APPLICATION NUMBER: 60/229,907
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/229,858
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
;   LENGTH: 297
;   TYPE: PRT
;   ORGANISM: Methylomonas 16a
US-09-941-947A-20

  Query Match           35.5%;  Score 516.5;  DB 9;  Length 297;
  Best Local Similarity  45.0%;  Pred. No. 1.6e-37;
  Matches  113;  Conservative   47;  Mismatches   88;  Indels    3;  Gaps    2;

Qy         23 KSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKSAIALEMIHTYSLIHD 82
              ::::   | ::| ||:  |||| ||:|   |  :|     :    | |:| || ||||||
Db         26 ENILPQTLHQAMRYSVLNGGKRTRPLLTYATGQALGLPENVLDAPACAVEFIHVYSLIHD 85

Qy         83 DLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSDDRLTDE--VKIKVLQR 140
              ||||||||| |||| | || | | |||||||||   |||::::|  :| :   ::|::
Db         86 DLPAMDNDDLRRGKPTCHKAYDEATAILAGDALQALAFEVLANDPGITVDAPARLKMITA 145

Qy        141 LSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFAV-MSAADIANVDDTT 199
              |: |||  |||||| :|: | |: : |  || :|  |||||:  :| ::|    ::|
Db        146 LTRASGSQGMVGGQAIDLGSVGRKLTLPELENMHIHKTGALIRASVNLAALSKPDLDTCV 205

Qy        200 KEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVSLLGKDGAEDKLT 259
               : |: |:  :|: ||:|||:||   | | |||  | |::|:| || :|||   ||: |
Db        206 AKKLDHYAKCIGLSFQVKDDILDIEADTATLGKTQGKDIDNDKPTYPALLGMAGAKQKAQ 265

Qy        260 YHRDAAVDELT 270
              : ||: ||
Db        266 ELHEQAVESLT 276


RESULT 9
```

US-09-934-903-14
; Sequence 14, Application US/09934903
; Patent No. US20020102690A1
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odom, J. Martin
; APPLICANT: Schenzle, Andreas J.
; APPLICANT: No. US20020102690Alton, Kelley C.
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Rouviere, Pierre
; APPLICANT: Picataggio, Stephen
; APPLICANT: Cheng, Qiong
; TITLE OF INVENTION: Genes Involved in Isoprenoid Compounds Production
; FILE REFERENCE: CL1646 US NA
; CURRENT APPLICATION NUMBER: US/09/934,903
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,907
; PRIOR FILING DATE: September 1, 2001
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Methylomonas 16a
; FEATURE:
; OTHER INFORMATION: Amino acid sequences encoded by ORF7
US-09-934-903-14

Query Match 35.5%; Score 516.5; DB 10; Length 297;
Best Local Similarity 45.0%; Pred. No. 1.6e-37;
Matches 113; Conservative 47; Mismatches 88; Indels 3; Gaps 2;

```
Qy         23 KSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKSAIALEMIHTYSLIHD 82
              ::::  | ::| ||:  |||| ||:|   |  :|     :    | |:| || ||||||
Db         26 ENILPQTLHQAMRYSVLNGGKRTRPLLTYATGQALGLPENVLDAPACAVEFIHVYSLIHD 85

Qy         83 DLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSDDRLTDE--VKIKVLQR 140
              ||||||||| |||| | || | |||||||||   |||:::::|  :| :   ::|::
Db         86 DLPAMDNDDLRRGKPTCHKAYDEATAILAGDALQALAFEVLANDPGITVDAPARLKMITA 145

Qy        141 LSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFAV-MSAADIANVDDTT 199
              |: |||  |||||| :|: | |: : |  || :|  |||||:  :| ::|    ::|
Db        146 LTRASGSQGMVGGQAIDLGSVGRKLTLPELENMHIHKTGALIRASVNLAALSKPDLDTCV 205

Qy        200 KEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVSLLGKDGAEDKLT 259
               : |: |:  :|: ||:|||:||   | | |||  | |::|:| || :|||  ||: |
Db        206 AKKLDHYAKCIGLSFQVKDDILDIEADTATLGKTQGKDIDNDKPTYPALLGMAGAKQKAQ 265

Qy        260 YHRDAAVDELT 270
                : ||: ||
Db        266 ELHEQAVESLT 276
```

RESULT 10
US-09-934-868-72
; Sequence 72, Application US/09934868
; Patent No. US20020137190A1
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odom, James M
; APPLICANT: Schenzle, Andreas J
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1596 US NA
; CURRENT APPLICATION NUMBER: US/09/934,868
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 72
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Methylomonas 16a
; FEATURE:
; OTHER INFORMATION: Amino acid sequences encoded by ORF7 - ISPa
US-09-934-868-72

Query Match 35.5%; Score 516.5; DB 10; Length 297;
Best Local Similarity 45.0%; Pred. No. 1.6e-37;
Matches 113; Conservative 47; Mismatches 88; Indels 3; Gaps 2;

```
Qy         23 KSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKSAIALEMIHTYSLIHD 82
              ::::  | ::| ||:  |||| ||:|   |  :|     :    | |:| || ||||||
Db         26 ENILPQTLHQAMRYSVLNGGKRTRPLLTYATGQALGLPENVLDAPACAVEFIHVYSLIHD 85

Qy         83 DLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSDDRLTDE--VKIKVLQR 140
              ||||||||| |||| | || | |||||||||   |||:::::|  :| :   ::|::
Db         86 DLPAMDNDDLRRGKPTCHKAYDEATAILAGDALQALAFEVLANDPGITVDAPARLKMITA 145

Qy        141 LSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFAV-MSAADIANVDDTT 199
              |: |||  |||||| :|: | |: : |  || :|  |||||:  :| ::|    ::|
Db        146 LTRASGSQGMVGGQAIDLGSVGRKLTLPELENMHIHKTGALIRASVNLAALSKPDLDTCV 205

Qy        200 KEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVSLLGKDGAEDKLT 259
               : |: |:  :|: ||:|||:||   | | |||  | |::|:| || :|||  ||: |
Db        206 AKKLDHYAKCIGLSFQVKDDILDIEADTATLGKTQGKDIDNDKPTYPALLGMAGAKQKAQ 265

Qy        260 YHRDAAVDELT 270
                : ||: ||
Db        266 ELHEQAVESLT 276
```

RESULT 11
US-10-108-915-26
; Sequence 26, Application US/10108915
; Patent No. US20020177204A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Shen, Jennie
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: Geranylgeranyl Pyrophosphate Synthases
; FILE REFERENCE: BB1286 US NA
; CURRENT APPLICATION NUMBER: US/10/108,915
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/452,238
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,592
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-December-02
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 26
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Glycine max
US-10-108-915-26

Query Match 34.7%; Score 504.5; DB 9; Length 367;
Best Local Similarity 44.1%; Pred. No. 2.5e-36;
Matches 116; Conservative 42; Mismatches 96; Indels 9; Gaps 3;

```
Qy         20 AINKSVMDT-------QLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKSAIALE 72
              |:||:: |         :: |:| ||| |||||:|||| :   : :  |    | :| |:|
Db         86 AVNKALDDAVSLREPQKIHEAMRYSLLAGGKRVRPVLCVAACELVGGEEATAMPAACAIE 145

Qy         73 MIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSDDRLTDE 132
              |||| |||||||| ||||| |||| |||||:||  |:|||||||  ||| |::  |
Db        146 MIHTMSLIHDDLPCMDNDDLRRGKPTNHKVFGEDVAVLAGDALLAFAFEHIAASTRGASP 205

Qy        133 VKI-KVLQRLSIASGHVGMVGGQMLDMQSEG-QPIDLETLEMIHKTKTGALLTFAVMSAA 190
              :| : :  |: : |  |:| ||::|: |||   :||| || ||  || ||| ||:  |
Db        206 GRIVRAIGELARSIGSEGLVAGQVVDINSEGLADVDLERLEFIHVHKTAALLEGAVVLGA 265

Qy        191 DIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVSLLG 250
               :    | |  :: ::|:::||: ||:||      :|||  | ||  :| || |||
Db        266 ILGGGTDDEVEKLRKFARYIGLLFQVVDDILDVTKSSQELGKTAGKDLVADKVTYPKLLG 325
```

```
Qy        251 KDGAEDKLTYHRDAAVDELTQID 273
              : ::         | |:|   |
Db        326 IEKSKVFAAKLNKDAQDQLVGFD 348


RESULT 12
US-09-815-242-11239
; Sequence 11239, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
   TITLE OF INVENTION:  Prokaryotes
   FILE REFERENCE: ELITRA.011A
   CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11239
;   LENGTH: 295
;   TYPE: PRT
;   ORGANISM: Haemophilus influenzae
US-09-815-242-11239

  Query Match             34.2%;  Score 497.5;  DB 10;  Length 295;
  Best Local Similarity   41.0%;  Pred. No. 7.5e-36;
  Matches  116;  Conservative   52;  Mismatches  100;  Indels   15;  Gaps    4;

Qy         16 ELSVAINKSVMDTQLE----------ESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGM 65
              ::   ||:  :: | |          |:| |:|  ||||:|| |:  |   |  | :
Db         10 QVQTRINR-FLEAQFEGIESHNAPLLEAMKYALLLGGKRVRPFLVYATGQMLGAEKQTLD 68

Qy         66 KSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISS 125
              :| |:| || |||||||||||||:|: |||  | |  : | |||||||| : |||:::
Db         69 YAAAAIEAIHAYSLIHDDLPAMDDDNLRRGHPTCHIQFDEATAILAGDALQSFAFEILTK 128

Qy        126 DDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFA 185
              :: | |: ::| |:  :|  ||  || ||: || : | |  ||:||: ||||||  |
Db        129 TPNISTEQKLALIQILAQGAGVQGMCLGQSLDLISEHKQISLSELELIHRNKTGALL-IA 187

Qy        186 VMSAADIAN---VDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNK 242
              :    | :    |   :: | |:  :|: ||::||:||  || |::||:||:||: :|
Db        188 ALKLGFICSPHFTDKRLEQSLTQYAEAIGLAFQVQDDILDIEGDSAEIGKQVGADLDLDK 247

Qy        243 STYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIV 285
              |||  |||  ||: |      :|: || :|      : | | :
Db        248 STYPKLLGLSGAKQKAQDLYQSALSELEKIPFDTTVRALAEFI 290


RESULT 13
US-10-108-915-22
; Sequence 22, Application US/10108915
; Patent No. US20020177204A1
; GENERAL INFORMATION:
;  APPLICANT: Cahoon, Rebecca E.
;  APPLICANT:  Shen, Jennie
;  APPLICANT:  Williams, Mark
;  TITLE OF INVENTION: Geranylgeranyl Pyrophosphate Synthases
;  FILE REFERENCE: BB1286 US NA
;  CURRENT APPLICATION NUMBER: US/10/108,915
;  CURRENT FILING DATE:  2002-03-28
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/452,238
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-01
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,592
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-December-02
;  NUMBER OF SEQ ID NOS: 46
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 22
;   LENGTH: 350
;   TYPE: PRT
;   ORGANISM: Glycine max
US-10-108-915-22

  Query Match             34.0%;  Score 494.5;  DB 9;  Length 350;
  Best Local Similarity   43.3%;  Pred. No. 1.7e-35;
  Matches  114;  Conservative   42;  Mismatches   98;  Indels    9;  Gaps    3;

Qy         20 AINKSVMDT-------QLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKSAIALE 72
              |:|| : |:        :: |:| ||| |||||:|||| :   : :       | :| |||
Db         69 AVNKGLDDSVSLREPKKIHEAMRYSLLAGGKRVRPVLCVAACELVGGHEATAMPAACALE 128

Qy         73 MIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSDDRLTDE 132
              |||| |||||||| ||||| |||| ||| |:||  |:|||||||  ||| |::  |
Db        129 MIHTMSLIHDDLPCMDNDDLRRGKPTNHTVFGEDVAVLAGDALLAFAFEHIAASTRGASA 188

Qy        133 VKI-KVLQRLSIASGHVGMVGGQMLDMQSEG-QPIDLETLEMIHKTKTGALLTFAVMSAA 190
              :| : :  |: : |  |:| ||::|: |||   : || || ||  || |||  ||:  |
Db        189 PRILRAIGELARSIGSEGLVAGQVVDINSEGLADVGLERLEFIHVHKTAALLEGAVVLGA 248

Qy        191 DIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVSLLG 250
              :    |  | |  :: ::|::||: ||:||      :|||  | ||  :| ||  |||
Db        249 ILGGGTDDEVEKLRKFARYIGLLFQVVDDILDVTKSSQELGKTAGKDLVADKVTYPKLLG 308

Qy        251 KDGAEDKLTYHRDAAVDELTQID 273
              : :::        | |:|   |
Db        309 IEKSKEFAAKLNKDAQDQLAGFD 331


RESULT 14
US-10-108-915-18
; Sequence 18, Application US/10108915
; Patent No. US20020177204A1
; GENERAL INFORMATION:
;  APPLICANT: Cahoon, Rebecca E.
;  APPLICANT:  Shen, Jennie
;  APPLICANT:  Williams, Mark
;  TITLE OF INVENTION: Geranylgeranyl Pyrophosphate Synthases
;  FILE REFERENCE: BB1286 US NA
;  CURRENT APPLICATION NUMBER: US/10/108,915
;  CURRENT FILING DATE:  2002-03-28
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/452,238
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-01
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,592
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-December-02
;  NUMBER OF SEQ ID NOS: 46
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 18
;   LENGTH: 369
;   TYPE: PRT
;   ORGANISM: Glycine max
US-10-108-915-18

  Query Match             33.6%;  Score 487.5;  DB 9;  Length 369;
  Best Local Similarity   44.9%;  Pred. No. 7.6e-35;
  Matches  111;  Conservative   43;  Mismatches   88;  Indels    5;  Gaps    4;
```

```
Qy         13 VNNELSVAINKSVMDT-QLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKSAIAL 71
              ||  |  ||  :: |  :: ::| ||| |||||:|||| :   : :       : :| |:
Db         89 VNQALDAAI--ALRDPHKIHQAMRYSLLAGGKRVRPVLCIAACELVGGTEATAIPAACAV 146

Qy         72 EMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFE-LISSDDRLT 130
              ||||| |||||||| ||||| |||| |||||||  |:|||||| ||| : :| : ::
Db        147 EMIHTMSLIHDDLPCMDNDDLRRGKPTNHKVYGEDVAVLAGDALLAFAFEHVAASTEGVS 206

Qy        131 DEVKIKVLQRLSIASGHVGMVGGQMLDMQSEG-QPIDLETLEMIHKTKTGALLTFAVMSA 189
                 :: :  |: : |  |:| ||::|: |||   : ||||| ||  || |||  ||:
Db        207 PSRVVRAIGELAKSIGTEGLVAGQVVDIDSEGVANVGLETLEFIHVHKTAALLEAAVVLG 266

Qy        190 ADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVSLL 249
              | :    |   | |  ::  :|::||: ||:||      :|||  | ||  :| ||  ||
Db        267 AIVGGGSDEEVEKLRKFARCIGLLFQVVDDILDVTKSSEELGKTAGKDLVADKVTYPKLL 326

Qy        250 GKDGAED 256
              | | :::
Db        327 GIDKSKE 333


RESULT 15
US-09-815-242-10069
; Sequence 10069, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10069
;   LENGTH: 299
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-09-815-242-10069

  Query Match             33.4%;  Score 485.5;  DB 10;  Length 299;
  Best Local Similarity   41.8%;  Pred. No. 8.6e-35;
  Matches  118;  Conservative   42;  Mismatches  117;  Indels    5;  Gaps    4;

Qy         10 IDEVNNELSVAINK-SVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKSA 68
              : : |  ||  |      :| : |:| |    ||||:|| |:  |                |
Db         11 VKQANQALSRFIAPLPFQNTPVVETMQYGALLGGKRLRPFLVYATGHMFGVSTNTLDAPA 70

Qy         69 IALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSDD- 127
                |:| || |||||||||||||:|| |||  | |  :||  |||||||| | || ::|  |
Db         71 AAVECIHAYSLIHDDLPAMDDDDLRRGLPTCHVKFGEANAILAGDALQTLAFSILSDADM 130

Qy        128 -RLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFAV 186
                 ::|  :| ::  |: |||  || ||| ||: :||: : |: || ||: |||||:  ||
Db        131 PEVSDRDRISMISELASASGIAGMCGGQALDLDAEGKHVPLDALERIHRHKTGALIRAAV 190

Qy        187 -MSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTY 245
                : |    :         |: |:  :|: ||::||:||  || | |||: |:| :  ||||
Db        191 RLGALSAGDKGRRALPVLDKYAESIGLAFQVQDDILDVVGDTATLGKRQGADQQLGKSTY 250

Qy        246 VSLLGKDGAEDKLTYHRDAAVDELTQIDEQ-FNTKHLLEIVD 286
                :||| : |  |      | |   | |: ||  :|  |  : |
Db        251 PALLGLEQARKKARDLIDDARQSLKQLAEQSLDTSALEALAD 292


Search completed: May 29, 2003, 12:08:17
Job time : 47 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 29, 2003, 11:56:10 ; Search time 19 Seconds
(without alignments)
1452.136 Million cell updates/sec

Title: US-09-925-637-64
Perfect score: 1453
Sequence: 1 MTNLPMNKLIDEVNNELSVA.....ELTQIDEQFNTKHLLEIVDL 287

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1444	99.4	293	2	A89932	hypothetical prote
2	702.5	48.3	293	2	AC1245	geranyltranstransf
3	686.5	47.2	293	2	AG1607	geranyltranstransf
4	654	45.0	297	2	JX0257	geranyltranstransf
5	638	43.9	294	2	E83997	geranyltranstransf
6	636.5	43.8	272	2	A69961	geranyltranstransf
7	549	37.8	291	2	G95139	geranyltranstransf
8	547	37.6	291	2	F98007	geranyltranstransf
9	540.5	37.2	285	2	A86732	geranyltranstransf
10	534.5	36.8	309	2	AE1833	geranylgeranyl dip
11	532	36.6	289	2	D97156	probable geranylge
12	530.5	36.5	300	2	T06969	farnesyltranstrans
13	526.5	36.2	302	2	S74538	geranylgeranyl pyr
14	497.5	34.2	295	2	C64123	geranyltranstransf
15	489.5	33.7	357	2	T09966	farnesyltranstrans
16	485.5	33.4	298	2	E81796	geranyltranstransf
17	485.5	33.4	299	2	JQ0665	geranyltranstransf
18	485.5	33.4	299	2	C90688	geranyltranstransf
19	485.5	33.4	299	2	G85538	geranyltranstransf
20	484	33.3	294	2	A82267	geranyltranstransf
21	479.5	33.0	369	2	S53722	farnesyltranstrans
22	479	33.0	366	2	T10452	farnesyltranstrans
23	478.5	32.9	316	2	T11021	farnesyltranstrans
24	478	32.9	371	2	F85434	geranylgeranyl pyr
25	475.5	32.7	306	2	AH0385	geranyltranstransf
26	474.5	32.7	295	2	F83139	geranyltranstransf
27	473	32.6	299	2	AG0554	geranyltranstransf
28	470.5	32.4	294	2	D87505	geranyltranstransf
29	469.5	32.3	259	2	F81217	geranyltranstransf
30	467	32.1	347	2	E84566	probable geranylge
31	457.5	31.5	304	2	F97685	geranyltransferase
32	457.5	31.5	335	2	AH2910	geranyltranstransf
33	438	30.1	327	2	D64407	bifunctional short
34	436	30.0	372	2	G84566	probable geranylge
35	434.5	29.9	304	2	AI3285	geranyltranstransf
36	434	29.9	376	2	T02429	geranylgeranyl pyr
37	419	28.8	378	2	S71230	geranylgeranyl pyr
38	417	28.7	281	2	B81261	geranyltranstransf
39	405	27.9	291	2	D82778	geranyltranstransf
40	378	26.0	325	2	T45152	bifunctional short
41	376.5	25.9	302	2	A37802	crtE protein - Erw
42	375	25.8	282	2	B84984	geranyltranstransf
43	373.5	25.7	303	2	A64636	geranyltranstransf
44	372.5	25.6	303	2	G71878	geranyltranstransf
45	372	25.6	346	2	C75139	bifunctional short

ALIGNMENTS

RESULT 1
A89932
hypothetical protein ispA [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: A89932
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of meticillin-resistant Stapylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: A89932
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-293 <KUR>
A;Cross-references: GB:BA000018; PID:g13701320; PIDN:BAB42614.1; GSPDB:GN00149
A;Experimental source: strain N315
C;Genetics:
A;Gene: ispA

Query Match 99.4%; Score 1444; DB 2; Length 293;
Best Local Similarity 99.3%; Pred. No. 1.9e-99;
Matches 285; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MTNLPMNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTNLPMNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTE 60

Qy         61 YELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAF 120
              ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YELGVKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAF 120

Qy        121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGA 180

Qy        181 LLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLEN 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        181 LLTFAVMSAADIANVDDATKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLEN 240

Qy        241 NKSTYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NKSTYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287
```

RESULT 2
AC1245
geranyltranstransferase homolog lmo1363 [imported] - Listeria monocytogenes (strain E
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 17-May-2002
C;Accession: AC1245

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AC1245
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-293 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAC99441.1; PID:g16410779; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo1363
C;Superfamily: geranyltranstransferase

Query Match 48.3%; Score 702.5; DB 2; Length 293;
Best Local Similarity 52.9%; Pred. No. 1.2e-44;
Matches 148; Conservative 46; Mismatches 81; Indels 5; Gaps 2;

```
Qy          8 KLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKS 67
              |:|||    |   ||:  :: :|:|||||: ||||||||:|:  || :|     ||:|:
Db         13 KVIDE---SLFKEINERNIEPRLKESMLYSIQAGGKRIRPMLVFATLQALKVNPLLGVKT 69

Qy         68 AIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSDD 127
              | |||||||||||||||||||||||||| |||||:|: ||||||||||| || ::: ||
Db         70 ATALEMIHTYSLIHDDLPAMDNDDYRRGKWTNHKVFGDATAILAGDALLTLAFSILAEDD 129

Qy        128 RLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFAVM 187
              |: | :| :: ::| :|| |||||: |:::| : : || |  ||  ||| || :||
Db        130 NLSFETRIALINQISFSSGAEGMVGGQLADLEAENKQVTLEELSSIHARKTGELLIYAVT 189

Qy        188 SAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVS 247
              ||| ||  |    : |  :: ::|: ||| ||:||  ||| |:||| |:|   |||||
Db        190 SAAKIAEADPEQTKRLRIFAENIGIGFQISDDILDVIGDETKMGKKTGADAFLNKSTYPG 249

Qy        248 LLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287
              ||  |||:  |  |   |   |:  |  |: : ||:: ||
Db        250 LLTLDGAKRALNEHVTIAKSALSGHD--FDDEILLKLADL 287
```

RESULT 3
AG1607
geranyltranstransferase homolog lin1400 [imported] - Listeria innocua (strain Clip11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 17-May-2002
C;Accession: AG1607
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AG1607
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-293 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC96631.1; PID:g16413873; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin1400
C;Superfamily: geranyltranstransferase

Query Match 47.2%; Score 686.5; DB 2; Length 293;
Best Local Similarity 52.1%; Pred. No. 1.8e-43;
Matches 146; Conservative 45; Mismatches 84; Indels 5; Gaps 2;

```
Qy          8 KLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKS 67
              |::||    |   ||   :: :|:|||||: ||||||||:|:  || :|| |   |:|:
Db         13 KVLDE---SLFKEINMRNIEPKLKESMLYSVQAGGKRIRPMLVFATLQALNIEPMRGLKT 69

Qy         68 AIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSDD 127
              | |||||||||||||||||||:|||||| ||||:||: |||||||||| || ::: |:
Db         70 ATALEMIHTYSLIHDDLPAMDDDDYRRGKWTNHKIYGDATAILAGDALLTLAFSILAEDE 129

Qy        128 RLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFAVM 187
              |: | :| :: ::| :|| |||||| ||::| : : || |  ||  ||| || |||
Db        130 NLSFETRIALINQISYSSGAEGMVGGQQADMEAENKQVTLEELASIHARKTGELLIFAVT 189

Qy        188 SAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVS 247
              ||| ||       : |  :: ::|: ||| ||:||  ||| |:||| | |   |||||
Db        190 SAAKIAEATPEQTKRLRIFAENIGIGFQISDDILDVIGDETKMGKKTGVDAFLNKSTYPG 249

Qy        248 LLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287
              ||  :||:  |  |   |   |:  |  |: : ||:: ||
Db        250 LLTLEGAKRALNEHVSIAKSALSGHD--FDDEILLKLADL 287
```

RESULT 4
JX0257
geranyltranstransferase (EC 2.5.1.10) - Bacillus stearothermophilus
N;Alternate names: farnesyl-diphosphate synthase
C;Species: Bacillus stearothermophilus
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 20-Jun-2000
C;Accession: JX0257
R;Koyama, T.; Obata, S.; Osabe, M.; Takeshita, A.; Yokoyama, K.; Uchida, M.; Nishino,
J. Biochem. 113, 355-363, 1993
A;Title: Thermostable farnesyl diphosphate synthase of Bacillus stearothermophilus: m
A;Reference number: JX0257; MUID:93252758; PMID:8486607
A;Accession: JX0257
A;Molecule type: DNA
A;Residues: 1-297 <KOY>
A;Cross-references: GB:D13293; NID:g391609; PIDN:BAA02551.1; PID:g391610
C;Genetics:
A;Start codon: GTG
C;Superfamily: geranyltranstransferase
C;Keywords: transferase

Query Match 45.0%; Score 654; DB 2; Length 297;
Best Local Similarity 47.7%; Pred. No. 4.8e-41;
Matches 133; Conservative 55; Mismatches 85; Indels 6; Gaps 2;

```
Qy          1 MTNLPMNKLIDEVNNELSVAINKSVM----DTQLEESMLYSLNAGGKRIRPVLLLLTLDS 56
              |  | : : ::|    :  |::: :       :|:::| ||| ||||||||:||| |: :
Db          1 MAQLSVEQFLNEQKQAVETALSRYIERLEGPAKLKKAMAYSLEAGGKRIRPLLLLSTVRA 60

Qy         57 LNTEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALL 116
              |  :  :|:  | |:|||||||||||||:|||| |||| ||||:||  |||||| ||
Db         61 LGKDPAVGLPVACAIEMIHTYSLIHDDLPSMDNDDLRRGKPTNHKVFGEAMAILAGDGLL 120

Qy        117 TKAFELIS--SDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIH 174
              | ||:||:   |:|:  |::::::||: |:|  ||| ||  ||: ||: : |  || ||
Db        121 TYAFQLITEIDDERIPPSVRLRLIERLAKAAGPEGMVAGQAADMEGEGKTLTLSELEYIH 180

Qy        175 KTKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKV 234
              : ||| :| ::| : | |   |       |: :: |||: |||:||:||  | | |:|| |
Db        181 RHKTGKMLQYSVHAGALIGGADARQTRELDEFAAHLGLAFQIRDDILDIEGAEEKIGKPV 240

Qy        235 GSDLENNKSTYVSLLGKDGAEDKLTYHRDAAVDELTQID 273
              |||  |||:|| :||   ||::|| :| :||   |   |
Db        241 GSDQSNNKATYPALLSLAGAKEKLAFHIEAAQRHLRNAD 279
```

RESULT 5
E83997
geranyltranstransferase BH2781 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 17-May-2002
C;Accession: E83997
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a

```
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: E83997
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-294 <STO>
A;Cross-references: GB:AP001516; GB:BA000004; NID:g10175192; PIDN:BAB06500.1; GSPDB:GN0
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH2781
C;Superfamily: geranyltranstransferase

  Query Match            43.9%;  Score 638;  DB 2;  Length 294;
  Best Local Similarity  47.7%;  Pred. No. 7.1e-40;
  Matches  134;  Conservative   53;  Mismatches   92;  Indels    2;  Gaps    1;

Qy          6 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGM 65
              :::: | :   :   | :      |: |||||| |||||||| ||| |: |   :   |:
Db          9 LDEIKDIIEERMPAHIERLNSPDMLKNSMLYSLKAGGKRIRPALLLATMKSFQKDISQGI 68

Qy         66 KSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISS 125
             | |:||||||||||||||:||:|| |||| |||||:||  |||||||||| :||:::
Db         69 DLACAIEMIHTYSLIHDDLPSMDDDDIRRGKPTNHKVFGEAHAILAGDALLTYSFEIVAK 128

Qy        126 DDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFA 185
              :     : ::: |: |:| |||||||: |:: | | : :| || ||  ||||||:||
Db        129 MKGVDPAKTLCLIEELARAAGPEGMVGGQVADIEGENQKLTVEGLEYIHHHKTGALLSFA 188

Qy        186 VMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTY 245
              ::: | :|:  :   |::  :|  ||::||||||:||  ||:| :|| ||||  | ||||
Db        189 IVAGARLADASEQDIENIRRFSRELGLLFQIKDDILDVEGDQAAIGKPVGSDDGNQKSTY 248

Qy        246 VSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVD 286
              |||  :||::||  |   | : | :  | | : | |: |
Db        249 PSLLTLEGAKEKLHLHTLLAKEYLYSV--QMNHRLLEELTD 287


RESULT 6
A69961
geranyltranstransferase homolog yqiD - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C;Accession: A69961
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Detter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: A69961
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-272 <KUN>
A;Cross-references: GB:Z99116; GB:AL009126; NID:g2634723; PIDN:CAB14359.1; PID:g2634862
A;Experimental source: strain 168
C;Genetics:
A;Gene: yqiD
C;Superfamily: geranyltranstransferase

  Query Match            43.8%;  Score 636.5;  DB 2;  Length 272;
  Best Local Similarity  49.3%;  Pred. No. 8.3e-40;
  Matches  134;  Conservative   51;  Mismatches   82;  Indels    5;  Gaps    2;

Qy          1 MTNLPMNKLIDE---VNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSL 57
                |||   : | |    : |:|||   |  |   |::|||||| |||||:||:::|  |::
Db          1 MTNKLTSFLADRKKTIENQLSVYTEKLDMPDSLKKSMLYSLQAGGKRLRPLIVLAVLNAY 60

Qy         58 NTEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLT 117
                 : |:    |:|||||||||||||| ||:|| |||| |||||:|| ||:|||| |||
Db         61 GKSEKDGIPVGCAVEMIHTYSLIHDDLPCMDDDDLRRGKPTNHKVFGEATAVLAGDGLLT 120

Qy        118 KAFELISS--DDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHK 175
              ::|:||:|   | :: | :::::  |  |:|  ||||||: ||::  : : || || ||:
Db        121 ESFKLITSHVSDEVSAEKRLRLVNELISAAGTEGMVGGQVADMEAGNRQVTLEELESIHE 180

Qy        176 TKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVG 235
               ||  || | |:: | :|:  :   | | ::| |:|: |||:||:||  | | |:||:||
Db        181 RKTSKLLGFCVIAGAILADAPEEDIETLRTFSSHIGIGFQIRDDILDLEGSEEKIGKRVG 240

Qy        236 SDLENNKSTYVSLLGKDGAEDKLTYHRDAAVD 267
              ||  |:|||| |||  :||: ||  |   | |
Db        241 SDTTNDKSTYPSLLSLEGAKHKLDVHIKEAND 272


RESULT 7
G95139
geranyltranstransferase [imported] - Streptococcus pneumoniae (strain TIGR4)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C;Accession: G95139
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzappl
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: G95139
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-291 <KUR>
A;Cross-references: GB:AE005672; PIDN:AAK75312.1; PID:g14972686; GSPDB:GN00164; TIGR:
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP1205

  Query Match            37.8%;  Score 549;  DB 2;  Length 291;
  Best Local Similarity  46.2%;  Pred. No. 2.7e-33;
  Matches  121;  Conservative   50;  Mismatches   87;  Indels    4;  Gaps    3;

Qy         28 TQLEESMLYSLNAGGKRIRPVLLLLTLDSLN-TEYELGMKSAIALEMIHTYSLIHDDLPA 86
              : | ||:|||::|||||||| |||  |::|  |       : | |||||| ||||||||
Db         25 SSLRESVLYSIHAGGKRIRPFLLLEVLEALQVTIKPAHAQVATALEMIHTGSLIHDDLPA 84

Qy         87 MDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSDDRLTDEVKIKVLQRLSIASG 146
              ||:||||||:|||| :||  ||||||||   :: ||:  | |  ::|: ::  ||:|||
Db         85 MDDDDYRRGRLTNHKKFGEAMAILAGDALFLDSYALIAQAD-LPSQIKVDLIANLSLASG 143

Qy        147 HVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFAVMSAADIANVDDTTKEHLESY 206
               :||| ||:|||: | | : || |: ||  ||| || :   :|| || :    :  |::
Db        144 SLGMVAGQVLDMEGEHQHLSLEELQTIHANKTGKLLAYPFQAAAIIAELSPEMQVKLKTV 203

Qy        207 SYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVSLLGKDGAEDKLTYHRDAAV 266
              :|: ||::||:||      ::||    ||:  |||| :||| : :       | |
Db        204 GELIGLAFQVRDDVLDVTASFEEIGKTPQKDLQAEKSTYPALLGLEESIAFCNQTLDQAN 263

Qy        267 DELTQIDEQ--FNTKHLLEIVD 286
              ::| :| :|  | |: :: :|:
Db        264 EKLEEIAQQLPFETESIVSVVE 285


RESULT 8
F98007
geranyltranstransferase (EC 2.5.1.10) [imported] - Streptococcus pneumoniae (strain R
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C;Accession: F98007
```

R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: F98007
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-291 <KUR>
A;Cross-references: GB:AE007317; PIDN:AAK99890.1; PID:g15458710; GSPDB:GN00174
C;Genetics:
A;Gene: ispA
C;Keywords: transferase

Query Match 37.6%; Score 547; DB 2; Length 291;
Best Local Similarity 45.8%; Pred. No. 3.8e-33;
Matches 120; Conservative 50; Mismatches 88; Indels 4; Gaps 3;

```
Qy          28 TQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYE-LGMKSAIALEMIHTYSLIHDDLPA 86
Db          25 SSLRESVLYSIHAGGKRIRPFLLLEVLEALQVAIKPAHAQVATALEMIHTGSLIHDDLPA 84

Qy          87 MDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSDDRLTDEVKIKVLQRLSIASG 146
Db          85 MDDDDYRRGRLTNHKKFGEAMAILAGDALFLDPYALIAQAD-LPSQIKVDLIANLSLASG 143

Qy         147 HVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFAVMSAADIANVDDTTKEHLESY 206
Db         144 SLGMVAGQVLDMEGEHQHLSLEELQTIHANKTGKLLAYPFQAAAIIAELSPEMQVKLKTV 203

Qy         207 SYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVSLLGKDGAEDKLTYHRDAAV 266
Db         204 GELIGLAFQVRDDVLDVTASFEEIGKTPQKDLQAEKSTYPALLGLEESIAFCNQTLDQAN 263

Qy         267 DELTQIDEQ--FNTKHLLEIVD 286
Db         264 EKLEEIAQQVPFETESIVSVVE 285
```

RESULT 9
A86732
geranyltranstransferase (EC 2.5.1.10) [imported] - Lactococcus lactis subsp. lactis (str
N;Alternate names: farnesyl diphosphate synthase
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C;Accession: A86732
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrli
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: A86732
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-285 <STO>
A;Cross-references: GB:AE005176; PID:g12723785; PIDN:AAK04955.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: ispA
C;Keywords: transferase

Query Match 37.2%; Score 540.5; DB 2; Length 285;
Best Local Similarity 44.8%; Pred. No. 1.1e-32;
Matches 130; Conservative 42; Mismatches 97; Indels 21; Gaps 5;

```
Qy           6 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTE----- 60
Db           1 MDTKILKLEDFLTEFYESAEFPTGLAESAKYSLLAGGKRIRPLLFLNLLEAFDLELSKAH 60

Qy          61 YELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAF 120
Db          61 YHV----AAALEMIHTGSLIHDDLPAMDNDDYRRGQLTNHKKFDEATAILAGDTLFFDPF 116

Qy         121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGA 180
Db         117 FILSTAD-LSAEIIVALTRELAFASGSYGMVAGQILDMAGEGKELTLAEIEQIHRLKTGR 175

Qy         181 LLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLEN 240
Db         176 LLTFPFVAAGIVAQKSTDEVEKLRQVGQILGLAFQIRDDILDVTATFAELGKTPGKDILE 235

Qy         241 NKSTYVSLLGKDGAEDKLTYH--------RDAAV---DELTQIDEQFNTK 279
Db         236 EKSTYVAHLGLEGAKKSLTGNLSEVKKLLTDLSVTDSSEIFKIIEQLEVK 285
```

RESULT 10
AE1833
geranylgeranyl diphosphate synthase [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp.
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C;Accession: AE1833
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE1833
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-309 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB77737.1; PID:g17135191; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr0213
C;Superfamily: geranyltranstransferase

Query Match 36.8%; Score 534.5; DB 2; Length 309;
Best Local Similarity 47.8%; Pred. No. 3.4e-32;
Matches 119; Conservative 35; Mismatches 92; Indels 3; Gaps 2;

```
Qy          29 QLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKSAIALEMIHTYSLIHDDLPAMD 88
Db          44 KIYESMRYSLLAGGKRLRPILCLATSEMMGGTIEIAMPTACAVEMIHTMSLIHDDLPAMD 103

Qy          89 NDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELIS--SDDRLTDEVKIKVLQRLSIASG 146
Db         104 NDDYRRGMLTNHKVYGEDIAILAGDGLLAYAFEFVAIRTPESVPRDRVLQVVARLGRALG 163

Qy         147 HVGMVGGQMLDMQSEGQ-PIDLETLEMIHKTKTGALLTFAVMSAADIANVDDTTKEHLES 205
Db         164 AAGLVGGQVVDLDSEGKSDTSLETLNFIHNHKTAALLEACVVCGGILAGASSENVQRLSR 223

Qy         206 YSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVSLLGKDGAEDKLTYHRDAA 265
Db         224 YSQNIGLAFQIIDDILDITSTQEQLGKTAGKDLLAKKVTYPSLWGIEQSRVKAQQLIEAA 283

Qy         266 VDELTQIDE 274
Db         284 CTELEPFGE 292
```

RESULT 11
D97156
probable geranylgeranyl pyrophosphate synthase [imported] - Clostridium acetobutylicu
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C;Accession: D97156
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: D97156

```
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-289 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK80039.1; PID:g15025068; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC2080

  Query Match            36.6%;  Score 532;  DB 2;  Length 289;
  Best Local Similarity  44.9%;  Pred. No. 4.8e-32;
  Matches  127;  Conservative   44;  Mismatches  108;  Indels    4;  Gaps    3;

Qy          7 NKLI-DEVNNELSVAI-NKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELG 64
              ||:|  ||   ||      |   : :: ||| |||||||||:||:||| :    | :|:
Db          3 NKVIKKEVEEYLSRYFEGKDNYNKRVYESMNYSLNAGGKRVRPLLLLASYAIYNKDYKEV 62

Qy         65 MKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELIS 124
              :  | |:|||||||||||||| ||||| |||: |||||::|   |:|||| || :|  ::
Db         63 IDIAAAIEMIHTYSLIHDDLPCMDNDDLRRGRPTNHKIFGYSIALLAGDGLLNEAMNIMF 122

Qy        125 SDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTF 184
                 |  :|    :| |:  ||:|||::|: |||: |: : |  :|| ||| |:
Db        123 KYCIGKGEEALKACLMISKAASSDGMIGGQVVDILSEGKKINEDELRYMHKKKTGELIKA 182

Qy        185 AVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKST 244
              ||:| | :        | |  |   ||: |||:||:||  ||   :||   |||||:| |
Db        183 AVVSGAILGGAPLHDVELLSQYGDKLGLAFQIEDDILDIIGDTKIMGKTSKSDLENDKCT 242

Qy        245 YVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287
              ||:| | |  :        |  :|: : :|  | ||: | || :|
Db        243 YVTLYGIDKCKSICRELTDECLDIIGKI--QGNTELLKEITEL 283


RESULT 12
T06969
farnesyltranstransferase (EC 2.5.1.29) - Cyanophora paradoxa cyanelle
C;Species: cyanelle Cyanophora paradoxa
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 24-Sep-1999
C;Accession: T06969
R;Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A.
submitted to the EMBL Data Library, July 1995
A;Description: Nucleotide sequence of the cyanelle genome from Cyanophora paradoxa.
A;Reference number: Z15840
A;Accession: T06969
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-300 <STI>
A;Cross-references: EMBL:U30821; NID:g1016083; PIDN:AAA81312.1; PID:g1016225
A;Experimental source: strain Pringsheim LB555
C;Genetics:
A;Gene: crtE
A;Genome: cyanelle
C;Superfamily: geranyltranstransferase
C;Keywords: carotenoid biosynthesis; chromoplast; cyanelle; transferase

  Query Match            36.5%;  Score 530.5;  DB 2;  Length 300;
  Best Local Similarity  43.0%;  Pred. No. 6.5e-32;
  Matches  120;  Conservative   51;  Mismatches  103;  Indels    5;  Gaps    3;

Qy          2 TNLPMNKLIDEVNNELSVAINKSV---MDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLN 58
              ||  :   : |    :   :|||:     | : :|: |||:||||||||:| | : :
Db          6 TNFNLLTYLYERKEIVEDTLNKSIPRGNPTFIYDSIRYSLSAGGKRIRPILCLASCELAG 65

Qy         59 TEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTK 118
                 |: : :| ||||||| |||||||||||||| ||||| ||| :|||  |||||||||
Db         66 GTMEMALPTACALEMIHTMSLIHDDLPAMDNDSYRRGKPTNHIIYGEDLAILAGDALLAY 125

Qy        119 AFELISSDDR-LTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQ-PIDLETLEMIHKT 176
              ||| |::  : :  :: :||: ::: :    |:||||::|: |||:    ||||  ||
Db        126 AFEFIATQTKNVPADLIVKVIAQVAHSVTTSGLVGGQIIDLSSEGKSDTTLETLNFIHIH 185

Qy        177 KTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGS 236
              ||||||  ||:| | :|   :        |: ::|: ||| ||:||    | |||| :|
Db        186 KTGALLEAAVLSGALLAGAKEKDMNRFLRYAQNIGLAFQIIDDVLDIISTEEKLGKSIGK 245

Qy        237 DLENNKSTYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQ 275
              ||:  |:|| | | : :  :      : | :|:   | :
Db        246 DLKTQKATYPSFWGVEESIKQAELLVEEAKEEILYFDNK 284


RESULT 13
S74538
geranylgeranyl pyrophosphate synthase - Synechocystis sp. (strain PCC 6803)
N;Alternate names: hypothetical protein slr0739
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S74538
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S74538
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-302 <KAN>
A;Cross-references: EMBL:D90899; GB:AB001339; NID:g1651650; PIDN:BAA16690.1; PID:g165
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Superfamily: geranyltranstransferase

  Query Match            36.2%;  Score 526.5;  DB 2;  Length 302;
  Best Local Similarity  44.2%;  Pred. No. 1.3e-31;
  Matches  117;  Conservative   49;  Mismatches   94;  Indels    5;  Gaps    3;

Qy          4 LPMNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYEL 63
              | : | : |    : |:|| :      :: |:| ||| |||||:||:| :    :    :  |
Db         16 LQVKKGVVEAALDSSLAIAR---PEKIYEAMRYSLLAGGKRLRPILCITACELCGGDEAL 72

Qy         64 GMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELI 123
              : :| ||||||| ||||||||:|||||:|||| |||||||  |||||| ||  ||| :
Db         73 ALPTACALEMIHTMSLIHDDLPSMDNDDFRRGKPTNHKVYGEDIAILAGDGLLAYAFEYV 132

Qy        124 SSDDRLTD-EVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQ-PIDLETLEMIHKTKTGAL 181
                 :    | :  ::|: ||    |  |:||||:||::|||:  |  |||  ||  |||||
Db        133 VTHTPQADPQALLQVIARLGRTVGAAGLVGGQVLDLESEGRTDITPETLTFIHTHKTGAL 192

Qy        182 LTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENN 241
              |  :|:: | :|      :: |  |: ::|: ||: ||:||    : :|||  | |::
Db        193 LEASVLTGAILAGATGEQQQRLARYAQNIGLAFQVVDDILDITATQEELGKTAGKDVKAQ 252

Qy        242 KSTYVSLLGKDGAEDKLTYHRDAAV 266
              |:|| |||| : :  :      | |:
Db        253 KATYPSLLGLEASRAQAQSLIDQAI 277


RESULT 14
C64123
geranyltranstransferase (EC 2.5.1.10) - Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 29-Sep-1999
C;Accession: C64123
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: C64123
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-295 <TIGR>
A;Cross-references: GB:U32822; GB:L42023; NID:g1574265; PIDN:AAC23087.1; PID:g1574277
C;Genetics:
```

```
A;Gene: ispA
C;Function:
A;Description: catalyzes condensation of isopentenyl diphosphate with dimethylallyl diph
A;Pathway: isoprenoid biosynthesis
C;Superfamily: dimethylallyltranstransferase
C;Keywords: isoprenoid biosynthesis; transferase
F;85-98/Region: aspartate-rich
F;223-227/Region: DDXXD motif

  Query Match            34.2%;  Score 497.5;  DB 2;  Length 295;
  Best Local Similarity  41.0%;  Pred. No. 1.7e-29;
  Matches  116;  Conservative   52;  Mismatches  100;  Indels   15;  Gaps    4;

Qy         16 ELSVAINKSVMDTQLE----------ESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGM 65
              ::   ||:. :: | |          |:| |:|  ||||:|| |:  |   |  | :
Db         10 QVQTRINR-FLEAQFEGIESHNAPLLEAMKYALLLGGKRVRPFLVYATGQMLGAEKQTLD 68

Qy         66 KSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISS 125
              :| |:| || |||||||||||:|: |||  | |  : | |||||||| : |||:::
Db         69 YAAAAIEAIHAYSLIHDDLPAMDDDNLRRGHPTCHIQFDEATAILAGDALQSFAFEILTK 128

Qy        126 DDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFA 185
              :: | |: ::| |:  :|  ||  || ||: || : | |  ||:||: ||||||  |
Db        129 TPNISTEQKLALIQILAQGAGVQGMCLGQSLDLISEHKQISLSELELIHRNKTGALL-IA 187

Qy        186 VMSAADIAN---VDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNK 242
               :   | :   |  :: | |:  :|: ||::||:||  || |::||:||:||: :|
Db        188 ALKLGFICSPHFTDKRLEQSLTQYAEAIGLAFQVQDDILDIEGDSAEIGKQVGADLDLDK 247

Qy        243 STYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIV 285
              |||  ||| ||: |     :|: || :|  .   : | | :
Db        248 STYPKLLGLSGAKQKAQDLYQSALSELEKIPFDTTVRALAEFI 290


RESULT 15
T09966
farnesyltranstransferase (EC 2.5.1.29) - Madagascar periwinkle
N;Alternate names: geranylgeranyl pyrophosphate synthase
C;Species: Catharanthus roseus (Madagascar periwinkle)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: T09966
R;Bantignies, B.; Liboz, T.; Ambid, C.
Plant Physiol. 110, 336, 1996
A;Title: Nucleotide sequence of a Catharanthus roseus geranylgeranyl pyrophosphate gene.
A;Reference number: Z16908
A;Accession: T09966
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-357 <BAN>
A;Cross-references: EMBL:X92893; NID:g1063275; PID:g1063276
A;Experimental source: cv. c20
C;Function:
A;Description: catalyzes the sequential addition of three molecules isopentenyl pyrophos
A;Pathway: isoprenoid biosynthesis
C;Superfamily: geranyltranstransferase
C;Keywords: transferase

  Query Match            33.7%;  Score 489.5;  DB 2;  Length 357;
  Best Local Similarity  41.9%;  Pred. No. 8.8e-29;
  Matches  111;  Conservative   48;  Mismatches   97;  Indels    9;  Gaps    3;

Qy         20 AINKSVMDT-------QLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKSAIALE 72
              ::||:: |        :: ||| ||| |||||:||:| :   :        : | || |:|
Db         76 SVNKALEDAVLVREPLKIHESMRYSLLAGGKRVRPMLCIAACELFGGTESVAMPSACAVE 135

Qy         73 MIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSDDR-LTD 131
              |||| ||:||||| ||||| |||| |||||:||  |:|||||||  ||| |::  : ::
Db        136 MIHTMSLMHDDLPCMDNDDLRRGKPTNHKVFGEDVAVLAGDALLAFAFEHIATATKGVSS 195

Qy        132 EVKIKVLQRLSIASGHVGMVGGQMLDMQSEG-QPIDLETLEMIHKTKTGALLTFAVMSAA 190
              |  ::|:  |:   |  |:| ||::|: |||   : || || ||  || |||  :|:  |
Db        196 ERIVRVVGELAKCIGSEGLVAGQVVDVCSEGIADVGLEHLEFIHIHKTAALLEGSVVLGA 255

Qy        191 DIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVSLLG 250
               :   :|      |  ::  :|::||: ||:||      :||| | ||  :| || |||
Db        256 IVGGANDEQISKLRKFARCIGLLFQVVDDILDVTKSSQELGKTAGKDLVADKVTYPKLLG 315

Qy        251 KDGAEDKLTYHRDAAVDELTQIDEQ 275
              | : :        | ::| : | :
Db        316 IDKSREFAEKLNREAQEQLAEFDPE 340


Search completed: May 29, 2003, 12:00:16
Job time : 21 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 29, 2003, 11:52:16 ; Search time 11 Seconds
(without alignments)
1082.155 Million cell updates/sec

Title: US-09-925-637-64
Perfect score: 1453
Sequence: 1 MTNLPMNKLIDEVNNELSVA.....ELTQIDEQFNTKHLLEIVDL 287

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	654	45.0	297	1	ISPA_BACST	Q08291	bacillus st
2	649	44.7	291	1	ISPA_MICLU	O66126	micrococcus
3	636.5	43.8	272	1	ISPA_BACSU	P54383	bacillus su
4	530.5	36.5	300	1	CRTE_CYAPA	P48368	cyanophora
5	497.5	34.2	295	1	ISPA_HAEIN	P45204	haemophilus
6	489.5	33.7	357	1	GGPP_CATRO	Q42698	catharanthu
7	485.5	33.4	299	1	ISPA_ECOLI	P22939	escherichia
8	479.5	33.0	369	1	GGPP_CAPAN	P80042	capsicum an
9	479	33.0	366	1	GGPP_SINAL	Q43133	sinapis alb
10	478	32.9	371	1	GGPP_ARATH	P34802	arabidopsis
11	438	30.1	327	1	IDSA_METJA	Q58270	methanococc
12	376.5	25.9	302	1	CRTE_PANAN	P21684	pantoea ana
13	376.5	25.9	324	1	IDSA_METTM	Q53479	methanobact
14	375	25.8	282	1	ISPA_BUCAI	P57537	buchnera ap
15	352	24.2	332	1	ISPA_RHISN	P55539	rhizobium s
16	346	23.8	307	1	CRTE_ERWHE	P22873	erwinia her
17	345.5	23.8	325	1	IDSA_METTH	O26156	methanobact
18	338	23.3	332	1	GGPP_SULSO	P95999	sulfolobus
19	334	23.0	332	1	ISPA_BRAJA	Q45220	bradyrhizob
20	330.5	22.7	262	1	ISPA_AQUAE	O66952	aquifex aeo
21	330	22.7	330	1	GGPP_SULAC	P39464	sulfolobus
22	320.5	22.1	289	1	CRTE_RHOCA	P17060	rhodobacter
23	311	21.4	323	1	ISPB_ECOLI	P19641	escherichia
24	302.5	20.8	288	1	CRTE_RHOSH	P54976	rhodobacter
25	294	20.2	329	1	ISPB_HAEIN	P44916	haemophilus
26	273	18.8	359	1	GGPP_MYCTU	Q50727	m probable
27	260.5	17.9	348	1	HEP2_BACSU	P31114	bacillus su
28	254	17.5	323	1	PREA_PORPU	P51268	porphyra pu
29	250	17.2	320	1	HEP2_BACST	P55785	bacillus st
30	248.5	17.1	323	1	PREA_CYAPA	P31171	cyanophora
31	245.5	16.9	323	1	PREA_SYNY3	P72580	synechocyst
32	211	14.5	323	1	PREA_CYACA	Q9tls1	cyanidium c
33	201.5	13.9	378	1	DPS_SCHPO	O43091	schizosacch
34	197.5	13.6	473	1	COQ1_YEAST	P18900	saccharomyc
35	173.5	11.9	352	1	FPPS_YEAST	P08524	saccharomyc
36	158.5	10.9	347	1	FPPS_SCHPO	O14230	schizosacch
37	156.5	10.8	367	1	FPPS_CHICK	P08836	gallus gall
38	149.5	10.3	353	1	FPPS_RAT	P05369	r farnesyl
39	149	10.3	349	1	FPPS_KLULA	P49349	kluyveromyc
40	147.5	10.2	300	1	GGPP_HUMAN	O95749	h geranylge
41	147.5	10.2	343	1	FPPS_ARTAN	P49350	artemisia a
42	141.5	9.7	294	1	GGPP_BOVIN	P56966	b geranylge
43	141.5	9.7	384	1	FPP1_ARATH	Q09152	arabidopsis
44	141	9.7	300	1	GGPP_MOUSE	Q9wtn0	m geranylge
45	141	9.7	353	1	FPPS_HUMAN	P14324	homo sapien

ALIGNMENTS

RESULT 1
ISPA_BACST
ID ISPA_BACST STANDARD; PRT; 297 AA.
AC Q08291; Q53435; Q53436; Q53437; Q53438;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Geranyltranstransferase (EC 2.5.1.10) (Farnesyl-diphosphate synthase)
DE (FPP synthase).
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillales; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=ATCC 10149;
RX MEDLINE=93252758; PubMed=8486607;
RA Koyama T., Obata S., Osabe M., Takeshita A., Yokoyama K.,
RA Uchida M., Nishino T., Ogura K.;
RT "Thermostable farnesyl diphosphate synthase of Bacillus
RT stearothermophilus: molecular cloning, sequence determination,
RT overproduction, and purification.";
RL J. Biochem. 113:355-363(1993).
RN [2]
RP MUTAGENESIS OF CYSTEINE RESIDUES.
RX MEDLINE=95001990; PubMed=7918490;
RA Koyama T., Obata S., Saito K., Takeshita-Koike A., Ogura K.;
RT "Structural and functional roles of the cysteine residues of Bacillus
RT stearothermophilus farnesyl diphosphate synthase.";
RL Biochemistry 33:12644-12648(1994).
CC -!- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
CC = diphosphate + trans,trans-farnesyl diphosphate.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
CC ----------

CC ----------
DR EMBL; D13293; BAA02551.1; -.
DR EMBL; S72629; AAB32272.1; -.
DR EMBL; S72630; AAB32273.2; ALT_SEQ.
DR EMBL; S72633; AAB32274.1; -.
DR EMBL; S72635; AAB32275.2; ALT_SEQ.
DR PIR; JX0257; JX0257.
DR InterPro; IPR000092; Polyprenyl_synt.
DR Pfam; PF00348; polyprenyl_synt; 1.
DR PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
DR PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
KW Transferase; Isoprene biosynthesis.
FT MUTAGEN 73 73 C->F,S: NO LOSS OF ACTIVITY.
FT MUTAGEN 289 289 C->F,S: NO LOSS OF ACTIVITY.
SQ SEQUENCE 297 AA; 32310 MW; 0F921C3F029EEBB6 CRC64;

```
  Query Match             45.0%;  Score 654;  DB 1;  Length 297;
  Best Local Similarity   47.7%;  Pred. No. 3.9e-41;
  Matches  133;  Conservative   55;  Mismatches   85;  Indels    6;  Gaps   2;

Qy          1 MTNLPMNKLIDEVNNELSVAINKSVM----DTQLEESMLYSLNAGGKRIRPVLLLLTLDS 56
              |  | : : ::|    :  |::: :      :|:::| ||| ||||||||:||| |: :
Db          1 MAQLSVEQFLNEQKQAVETALSRYIERLEGPAKLKKAMAYSLEAGGKRIRPLLLLSTVRA 60

Qy         57 LNTEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALL 116
              |  :  :|: | |:|||||||||||||:||||| |||| ||||||:||  |||||| ||
Db         61 LGKDPAVGLPVACAIEMIHTYSLIHDDLPSMDNDDLRRGKPTNHKVFGEAMAILAGDGLL 120

Qy        117 TKAFELIS--SDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIH 174
              | ||:||:   |:|:   |:::::||: |:|  ||| ||  ||: ||: : |  || ||
Db        121 TYAFQLITEIDDERIPPSVRLRLIERLAKAAGPEGMVAGQAADMEGEGKTLTLSELEYIH 180

Qy        175 KTKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKV 234
              : ||| :| ::| : | |   |      |: :: ||| |||:||:||  | | |:|| |
Db        181 RHKTGKMLQYSVHAGALIGGADARQTRELDEFAAHLGLAFQIRDDILDIEGAEEKIGKPV 240

Qy        235 GSDLENNKSTYVSLLGKDGAEDKLTYHRDAAVDELTQID 273
              |||  |||:|| :||   ||::|| :| :||   |   |
Db        241 GSDQSNNKATYPALLSLAGAKEKLAFHIEAAQRHLRNAD 279


RESULT 2
ISPA_MICLU
ID   ISPA_MICLU     STANDARD;      PRT;   291 AA.
AC   O66126;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Geranyltranstransferase (EC 2.5.1.10) (Farnesyl-diphosphate synthase)
DE   (FPP synthase).
GN   FPS.
OS   Micrococcus luteus (Micrococcus lysodeikticus).
OC   Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC   Actinomycetales; Micrococcineae; Micrococcaceae; Micrococcus.
OX   NCBI_TaxID=1270;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B-P 26;
RX   MEDLINE=98175686; PubMed=9515931;
RA   Shimizu N., Koyama T., Ogura K.;
RT   "Molecular cloning, expression, and characterization of the genes
RT   encoding the two essential protein components of Micrococcus luteus
RT   B-P 26 hexaprenyl diphosphate synthase.";
RL   J. Bacteriol. 180:1578-1581(1998).
CC   -!- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
CC       = diphosphate + trans,trans-farnesyl diphosphate.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB003187; BAA25265.1; -.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
KW   Transferase; Isoprene biosynthesis.
SQ   SEQUENCE   291 AA;  32362 MW;  797125AB71E5674A CRC64;

  Query Match             44.7%;  Score 649;  DB 1;  Length 291;
  Best Local Similarity   52.9%;  Pred. No. 8.9e-41;
  Matches  146;  Conservative   40;  Mismatches   78;  Indels   12;  Gaps   4;

Qy         13 VNNELSVAINKSVMD------TQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMK 66
              :| :    ||:|:::      ::| |:: |||:||||||||:|:| |||||      |:
Db          8 MNRDFLNLINESLLNKYHPAQSRLHEAINYSLSAGGKRIRPLLVLTTLDSLGGNAHDGLP 67

Qy         67 SAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSD 126
              |||||||||||||||||||||||||||||||| : | |||||||||| ||: |  :
Db         68 FGIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKRFDEATAILAGDALLTDAFQCI-LN 126

Qy        127 DRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFAV 186
               :|  |:|: ::  || |||  ||| ||||||| | : : |  || ||  ||| |:  |:
Db        127 TQLNAEIKLSLINLLSTASGSNGMVYGQMLDMQGEHKTLTLNELERIHIHKTGELIRAAI 186

Qy        187 MSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYV 246
              :||  | | :|   | |      ::|:||||||:||  |    :|| ||||| |:|||||
Db        187 VSAGIIMNFNDAQIEQLNIIGKNVGLMFQIKDDILDVEGSFENIGKTVGSDLNNDKSTYV 246

Qy        247 SLLGKDGAE----DKLTYHRDAAVDELTQIDEQFNT 278
              |||| : ::    ||||   | |:  |  |::   |
Db        247 SLLGLEASKQLLNDKLTETYD-ALKTLQPINDNLKT 281


RESULT 3
ISPA_BACSU
ID   ISPA_BACSU     STANDARD;      PRT;   272 AA.
AC   P54383;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Geranyltranstransferase (EC 2.5.1.10) (Farnesyl-diphosphate synthase)
DE   (FPP synthase).
GN   YQID.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168 / JH642;
RX   MEDLINE=97124195; PubMed=8969508;
RA   Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
RA   Kobayashi Y.;
RT   "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of
RT   the Bacillus subtilis genome containing the skin element and many
RT   sporulation genes.";
RL   Microbiology 142:3103-3111(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
```

```
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the Gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
CC   -!- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
CC       = diphosphate + trans,trans-farnesyl diphosphate.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D84432; BAA12575.1; -.
DR   EMBL; Z99116; CAB14359.1; -.
DR   SubtiList; BG11714; yqiD.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
KW   Transferase; Isoprene biosynthesis; Complete proteome.
SQ   SEQUENCE   272 AA;  29740 MW;  55A26E9A9895D67C CRC64;

  Query Match           43.8%;  Score 636.5;  DB 1;  Length 272;
  Best Local Similarity  49.3%;  Pred. No. 6.7e-40;
  Matches  134;  Conservative   51;  Mismatches   82;  Indels    5;  Gaps    2;

Qy          1 MTNLPMNKLIDE---VNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSL 57
              |||   : | |     : |:|||   |   |    |::|||||| |||||:||:::|  |::
Db          1 MTNKLTSFLADRKKTIENQLSVYTEKLDMPDSLKKSMLYSLQAGGKRLRPLIVLAVLNAY 60

Qy         58 NTEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLT 117
                : |:    |:|||||||||||||| ||:|| |||| |||||:|| ||:|||| |||
Db         61 GKSEKDGIPVGCAVEMIHTYSLIHDDLPCMDDDDLRRGKPTNHKVFGEATAVLAGDGLLT 120

Qy        118 KAFELISS--DDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHK 175
              ::|:||:|   | :: | :::::  |  |:|  ||||||: ||::  : : || || ||:
Db        121 ESFKLITSHVSDEVSAEKRLRLVNELISAAGTEGMVGGQVADMEAGNRQVTLEELESIHE 180

Qy        176 TKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVG 235
               ||  || | |:: | :|:  :   | | ::| |:|: |||:||:||  | | |:||:||
Db        181 RKTSKLLGFCVIAGAILADAPEEDIETLRTFSSHIGIGFQIRDDILDLEGSEEKIGKRVG 240

Qy        236 SDLENNKSTYVSLLGKDGAEDKLTYHRDAAVD 267
               ||  |:|||| |||  :||: ||  |   | |
Db        241 SDTTNDKSTYPSLLSLEGAKHKLDVHIKEAND 272


RESULT 4
CRTE_CYAPA
ID   CRTE_CYAPA     STANDARD;      PRT;   300 AA.
AC   P48368;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Geranylgeranyl pyrophosphate synthetase homolog (EC 2.5.1.29) (GGPP
DE   synthetase) (Farnesyltranstransferase).
GN   CRTE.
OS   Cyanophora paradoxa.
OG   Cyanelle.
OC   Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.
OX   NCBI_TaxID=2762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LB555 / Pringsheim;
RA   Stirewalt V.L., Michalowski C.B., Loeffelhardt W., Bohnert H.J.,
RA   Bryant D.A.;
RT   "Nucleotide sequence of the cyanelle DNA from Cyanophora paradoxa.";
RL   Plant Mol. Biol. Rep. 13:327-332(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LB555 / Pringsheim;
RA   Loeffelhardt W., Stirewalt V.L., Michalowski C.B., Annarella M.,
RA   Farley J.Y., Schluchter W.M., Chung S., Newmann-Spallart C.,
RA   Steiner J.M., Jakowitsch J., Bohnert H.J., Bryant D.A.;
RT   "The complete sequence of the cyanelle genome of Cyanophora paradoxa:
RT   the genetic complexity of a primitive plastid.";
RL   (In) Schenk H.E.A., Herrmann R., Jeon K.W., Mueller N.E.,
RL   Schwemmler W. (eds.);
RL   Eukaryotism and Symbiosis, pp.40-48, Springer-Verlag, Heidelberg
RL   (1997).
CC   -!- CATALYTIC ACTIVITY: Trans-trans-farnesyl diphosphate + isopentenyl
CC       diphosphate = diphosphate + geranylgeranyl diphosphate.
CC   -!- PATHWAY: Carotenoid and chlorophyll biosynthesis.
CC   -!- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U30821; AAA81312.1; -.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
KW   Photosynthesis; Chlorophyll biosynthesis; Carotenoid biosynthesis;
KW   Isoprene biosynthesis; Transferase; Cyanelle.
SQ   SEQUENCE   300 AA;  33008 MW;  05950D0E3BA04B34 CRC64;

  Query Match           36.5%;  Score 530.5;  DB 1;  Length 300;
  Best Local Similarity  43.0%;  Pred. No. 4.6e-32;
  Matches  120;  Conservative   51;  Mismatches  103;  Indels    5;  Gaps    3;

Qy          2 TNLPMNKLIDEVNNELSVAINKSV---MDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLN 58
              ||  :   : |    :   :|||:     | : :|: |||:|||||||:| | : :
Db          6 TNFNLLTYLYERKEIVEDTLNKSIPRGNPTFIYDSIRYSLSAGGKRIRPILCLASCELAG 65

Qy         59 TEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTK 118
                |: : :| ||||||| |||||||||||| ||||| ||| :|||  |||||||||
Db         66 GTMEMALPTACALEMIHTMSLIHDDLPAMDNDSYRRGKPTNHIIYGEDLAILAGDALLAY 125

Qy        119 AFELISSDDR-LTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQ-PIDLETLEMIHKT 176
              ||| |::  : :  :: :||: ::: :    |:||||::|: |||:    |||| ||
Db        126 AFEFIATQTKNVPADLIVKVIAQVAHSVTTSGLVGGQIIDLSSEGKSDTTLETLNFIHIH 185

Qy        177 KTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGS 236
              ||||||  ||:| | :|   :         |: ::|: ||| ||:||    | |||| :|
Db        186 KTGALLEAAVLSGALLAGAKEKDMNRFLRYAQNIGLAFQIIDDVLDIISTEEKLGKSIGK 245

Qy        237 DLENNKSTYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQ 275
              ||:  |:|| |  | : :  :      : | :|:   | :
Db        246 DLKTQKATYPSFWGVEESIKQAELLVEEAKEEILYFDNK 284


RESULT 5
ISPA_HAEIN
ID   ISPA_HAEIN     STANDARD;      PRT;   295 AA.
AC   P45204;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Geranyltranstransferase (EC 2.5.1.10) (Farnesyl-diphosphate synthase)
DE   (FPP synthase).
GN   ISPA OR HI1438.
OS   Haemophilus influenzae.
```

```
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Rd / KW20 / ATCC 51907;
RX   MEDLINE=95350630; PubMed=7542800;
RA   Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA   Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA   McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT   Rd.";
RL   Science 269:496-512(1995).
CC   -!- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
CC       = diphosphate + trans,trans-farnesyl diphosphate.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U32822; AAC23087.1; -.
DR   TIGR; HI1438; -.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
KW   Transferase; Isoprene biosynthesis; Complete proteome.
SQ   SEQUENCE   295 AA;  32407 MW;  A64DE650266E670F CRC64;

  Query Match           34.2%;  Score 497.5;  DB 1;  Length 295;
  Best Local Similarity  41.0%;  Pred. No. 1.2e-29;
  Matches  116;  Conservative   52;  Mismatches  100;  Indels   15;  Gaps    4;

Qy         16 ELSVAINKSVMDTQLE----------ESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGM 65
              ::   ||:  :: | |          |:| |:|  ||||:|| |:  |   |  | :
Db         10 QVQTRINR-FLEAQFEGIESHNAPLLEAMKYALLGGKRVRPFLVYATGQMLGAEKQTLD 68

Qy         66 KSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISS 125
               :| |:| || |||||||||||:|: |||  | |  : | |||||||| : |||:::
Db         69 YAAAAIEAIHAYSLIHDDLPAMDDDNLRRGHPTCHIQFDEATAILAGDALQSFAFEILTK 128

Qy        126 DDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFA 185
              :: | |: ::| |:  :|  ||  || ||: || : | |  ||:||: ||||||  |
Db        129 TPNISTEQKLALIQILAQGAGVQGMCLGQSLDLISEHKQISLSELELIHRNKTGALL-IA 187

Qy        186 VMSAADIAN---VDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNK 242
               :   | :    |  :: |  |:  :|: ||::||:||  || |::||:||:||: :|
Db        188 ALKLGFICSPHFTDKRLEQSLTQYAEAIGLAFQVQDDILDIEGDSAEIGKQVGADLDLDK 247

Qy        243 STYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIV 285
              |||  |||  ||: |      :|: || :|      : | | :
Db        248 STYPKLLGLSGAKQKAQDLYQSALSELEKIPFDTTVRALAEFI 290


RESULT 6
GGPP_CATRO
ID   GGPP_CATRO     STANDARD;      PRT;   357 AA.
AC   Q42698;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Geranylgeranyl pyrophosphate synthetase, chloroplast precursor (GGPP
DE   synthetase) (GGPS) [Includes: Dimethylallyltransferase (EC 2.5.1.1);
DE   Geranyltranstransferase (EC 2.5.1.10); Farnesyltranstransferase
DE   (EC 2.5.1.29)].
GN   GGPS1 OR GGC1.
OS   Catharanthus roseus (Rosy periwinkle) (Madagascar periwinkle).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Gentianales; Apocynaceae; Rauvolfioideae;
OC   Vinceae; Catharanthus.
OX   NCBI_TaxID=4058;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. G. Don c20;
RA   Bantignies B., Liboz T., Ambid C.;
RT   "Nucleotide sequence of a Catharanthus roseus geranylgeranyl
RT   pyrophosphate synthase gene.";
RL   (In) Plant Gene Register PGR95-119.
CC   -!- FUNCTION: CATALYZES THE TRANS-ADDITION OF THE THREE MOLECULES OF
CC       IPP ONTO DMAPP TO FORM GERANYLGERANYL PYROPHOSPHATE.
CC   -!- CATALYTIC ACTIVITY: Dimethylallyl diphosphate + isopentenyl
CC       diphosphate = diphosphate + geranyl diphosphate.
CC   -!- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
CC       = diphosphate + trans,trans-farnesyl diphosphate.
CC   -!- CATALYTIC ACTIVITY: Trans-trans-farnesyl diphosphate + isopentenyl
CC       diphosphate = diphosphate + geranylgeranyl diphosphate.
CC   -!- PATHWAY: FIRST COMMITTED STEP IN CAROTENOID BIOSYNTHESIS. KEY
CC       ENZYME IN PLANT TERPENOID BIOSYNTHESIS.
CC   -!- SUBCELLULAR LOCATION: Chloroplast.
CC   -!- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X92893; CAA63486.1; -.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
KW   Transferase; Isoprene biosynthesis; Carotenoid biosynthesis;
KW   Chloroplast; Transit peptide.
FT   TRANSIT       1     40       CHLOROPLAST (POTENTIAL).
FT   CHAIN        41    357       GERANYLGERANYL PYROPHOSPHATE SYNTHETASE.
SQ   SEQUENCE   357 AA;  38786 MW;  81C52FDEA1E06FA8 CRC64;

  Query Match           33.7%;  Score 489.5;  DB 1;  Length 357;
  Best Local Similarity  41.9%;  Pred. No. 5.9e-29;
  Matches  111;  Conservative   48;  Mismatches   97;  Indels    9;  Gaps    3;

Qy         20 AINKSVMDT-------QLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKSAIALE 72
              ::||:: |        :: ||| ||| |||||:||:| :   :       : | || |:|
Db         76 SVNKALEDAVLVREPLKIHESMRYSLLAGGKRVRPMLCIAACELFGGTESVAMPSACAVE 135

Qy         73 MIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSDDR-LTD 131
              |||| ||:|||| |||| |||| |||||:||  |:|||||  ||| |::  : ::
Db        136 MIHTMSLMHDDLPCMDNDDLRRGKPTNHKVFGEDVAVLAGDALLAFAFEHIATATKGVSS 195

Qy        132 EVKIKVLQRLSIASGHVGMVGGQMLDMQSEG-QPIDLETLEMIHKTKTGALLTFAVMSAA 190
              |  ::|:  |:   |  |:| ||::|: |||   : || || ||  || |||  :|:  |
Db        196 ERIVRVVGELAKCIGSEGLVAGQVVDVCSEGIADVGLEHLEFIHIHKTAALLEGSVVLGA 255

Qy        191 DIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVSLLG 250
               :   :|     |  ::  :|::||: ||:||      :|||  | ||  :| ||  |||
Db        256 IVGGANDEQISKLRKFARCIGLLFQVVDDILDVTKSSQELGKTAGKDLVADKVTYPKLLG 315

Qy        251 KDGAEDKLTYHRDAAVDELTQIDEQ 275
              | : :        | ::| : | :
Db        316 IDKSREFAEKLNREAQEQLAEFDPE 340
```

```
RESULT 7
ISPA_ECOLI
ID   ISPA_ECOLI     STANDARD;      PRT;   299 AA.
AC   P22939;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Geranyltranstransferase (EC 2.5.1.10) (Farnesyl-diphosphate synthase)
DE   (FPP synthase).
GN   ISPA OR B0421.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=91210228; PubMed=2089044;
RA   Fujisaki S., Hara H., Nishimura Y., Horiuchi K., Nishino T.;
RT   "Cloning and nucleotide sequence of the ispA gene responsible for
RT   farnesyl diphosphate synthase activity in Escherichia coli.";
RL   J. Biochem. 108:995-1000(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Roberts D., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,
RA   Duncan M., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
RA   Lew H., Lin D., Namath A., Oefner P., Schramm S., Davis R.W.;
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
CC       = diphosphate + trans,trans-farnesyl diphosphate.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00694; BAA00599.1; -.
DR   EMBL; AE000148; AAC73524.1; -.
DR   EMBL; U82664; AAB40177.1; -.
DR   PIR; JQ0665; JQ0665.
DR   SWISS-2DPAGE; P22939; COLI.
DR   EcoGene; EG10508; ispA.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
KW   Transferase; Isoprene biosynthesis; Complete proteome.
SQ   SEQUENCE   299 AA;  32159 MW;  15BADD5E135060CA CRC64;

  Query Match             33.4%;  Score 485.5;  DB 1;  Length 299;
  Best Local Similarity   41.8%;  Pred. No. 9.3e-29;
  Matches  118;  Conservative   42;  Mismatches  117;  Indels    5;  Gaps    4;

Qy         10 IDEVNNELSVAINK-SVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKSA 68
              : : |  ||  |      :| : |:| |    ||||:|| |:  |               |
Db         11 VKQANQALSRFIAPLPFQNTPVVETMQYGALLGGKRLRPFLVYATGHMFGVSTNTLDAPA 70

Qy         69 IALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSDD- 127
              |:| || |||||||||||:|| |||  | |  :||  |||||||| | || ::|  |
Db         71 AAVECIHAYSLIHDDLPAMDDDDLRRGLPTCHVKFGEANAILAGDALQTLAFSILSDADM 130

Qy        128 -RLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFAV 186
               ::|  :| ::  |: |||  || ||| ||: :||: : |: || ||: |||||:  ||
Db        131 PEVSDRDRISMISELASASGIAGMCGGQALDLDAEGKHVPLDALERIHRHKTGALIRAAV 190

Qy        187 -MSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTY 245
               : |   :          |: |:  :|: ||::||:||  || | |||: |:| :  ||||
Db        191 RLGALSAGDKGRRALPVLDKYAESIGLAFQVQDDILDVVGDTATLGKRQGADQQLGKSTY 250

Qy        246 VSLLGKDGAEDKLTYHRDAAVDELTQIDEQ-FNTKHLLEIVD 286
              :||| : |  |     | |   | |: ||  :|  |  : |
Db        251 PALLGLEQARKKARDLIDDARQSLKQLAEQSLDTSALEALAD 292


RESULT 8
GGPP_CAPAN
ID   GGPP_CAPAN     STANDARD;      PRT;   369 AA.
AC   P80042;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Geranylgeranyl pyrophosphate synthetase, chloroplast precursor (GGPP
DE   synthetase) (GGPS) [Includes: Dimethylallyltransferase (EC 2.5.1.1);
DE   Geranyltranstransferase (EC 2.5.1.10); Farnesyltranstransferase
DE   (EC 2.5.1.29)].
GN   GGPS1.
OS   Capsicum annuum (Bell pepper).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Capsicum.
OX   NCBI_TaxID=4072;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Lamuyo; TISSUE=Fruit;
RX   MEDLINE=93272043; PubMed=1303794;
RA   Kuntz M., Roemer S., Suire C., Hugueney P., Weil J.H., Schantz R.,
RA   Camara B.;
RT   "Identification of a cDNA for the plastid-located geranylgeranyl
RT   pyrophosphate synthase from Capsicum annuum: correlative increase in
RT   enzyme activity and transcript level during fruit ripening.";
RL   Plant J. 2:25-34(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Yolo Wonder;
RX   MEDLINE=95195169; PubMed=7888631;
RA   Badillo A., Steppuhn J., Deruere J., Camara B., Kuntz M.;
RT   "Structure of a functional geranylgeranyl pyrophosphate synthase gene
RT   from Capsicum annuum.";
RL   Plant Mol. Biol. 27:425-428(1995).
CC   -!- FUNCTION: CATALYZES THE TRANS-ADDITION OF THE THREE MOLECULES OF
CC       IPP ONTO DMAPP TO FORM GERANYLGERANYL PYROPHOSPHATE.
CC   -!- CATALYTIC ACTIVITY: Dimethylallyl diphosphate + isopentenyl
CC       diphosphate = diphosphate + geranyl diphosphate.
CC   -!- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
CC       = diphosphate + trans,trans-farnesyl diphosphate.
CC   -!- CATALYTIC ACTIVITY: Trans-trans-farnesyl diphosphate + isopentenyl
CC       diphosphate = diphosphate + geranylgeranyl diphosphate.
CC   -!- PATHWAY: FIRST COMMITTED STEP IN CAROTENOID BIOSYNTHESIS. KEY
CC       ENZYME IN PLANT TERPENOID BIOSYNTHESIS.
CC   -!- SUBUNIT: MONOMER.
CC   -!- SUBCELLULAR LOCATION: Chloroplast.
CC   -!- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; X80267; CAA56554.1; -.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
KW   Transferase; Isoprene biosynthesis; Carotenoid biosynthesis;
KW   Chloroplast; Transit peptide.
FT   TRANSIT       1      ?       CHLOROPLAST.
FT   CHAIN         ?    369       GERANYLGERANYL PYROPHOSPHATE SYNTHETASE.
SQ   SEQUENCE   369 AA;  40173 MW;  2D527F5E43A29C6F CRC64;

  Query Match             33.0%;  Score 479.5;  DB 1;  Length 369;
  Best Local Similarity   42.4%;  Pred. No. 3.4e-28;
  Matches  106;  Conservative   48;  Mismatches   87;  Indels    9;  Gaps    3;

Qy         16 ELSVAINKSVMDT-------QLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKSA 68
              | ::::||:: :         : |:| ||| |||||:||:| |   : :    |  | :|
Db         84 EKAISVNKALDEAIIVKEPHVIHEAMRYSLLAGGKRVRPMLCLAACELVGGNQENAMAAA 143

Qy         69 IALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFE-LISSDD 127
              |:|||| |||||||| ||||| |||| ||||:||| |:||||:|| ||| :::|
Db        144 CAVEMIHTMSLIHDDLPCMDNDDLRRGKPTNHKIYGEDVAVLAGDSLLAFAFEHIVNSTA 203

Qy        128 RLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQ-PIDLETLEMIHKTKTGALLTFAV 186
              :|    :  :  |: : |  |:| ||: |::  |   : ||||| ||  || |||  :|
Db        204 GVTPSRIVGAVAELAKSIGTEGLVAGQVADIKCTGNASVSLETLEFIHVHKTAALLESSV 263

Qy        187 MSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYV 246
              :  | :    :   | |  ::  :|::||: ||:||      :|||  | ||  :|:||
Db        264 VLGAILGGGTNVEVEKLRRFARCIGLLFQVVDDILDVTKSSEELGKTAGKDLVVDKTTYP 323

Qy        247 SLLGKDGAED 256
              ||| : |::
Db        324 KLLGLEKAKE 333


RESULT 9
GGPP_SINAL
ID   GGPP_SINAL     STANDARD;      PRT;   366 AA.
AC   Q43133;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Geranylgeranyl pyrophosphate synthetase, chloroplast precursor (GGPP
DE   synthetase) (GGPS) [Includes: Dimethylallyltransferase (EC 2.5.1.1);
     Geranyltranstransferase (EC 2.5.1.10); Farnesyltranstransferase
     (EC 2.5.1.29)].
     GGPS1 OR GGPS.
OS   Sinapis alba (White mustard) (Brassica hirta).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Sinapis.
OX   NCBI_TaxID=3728;
RN   [1]
RP   SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
RX   MEDLINE=97433278; PubMed=9288918;
RA   Bonk M., Hoffmann B., von Lintig J., Schledz M., Al-Babili S.,
RA   Hobeika E., Kleinig H., Beyer P.;
RT   "Chloroplast import of four carotenoid biosynthetic enzymes in vitro
RT   reveals differential fates prior to membrane binding and oligomeric
RT   assembly.";
RL   Eur. J. Biochem. 247:942-950(1997).
CC   -!- FUNCTION: CATALYZES THE TRANS-ADDITION OF THE THREE MOLECULES OF
CC       IPP ONTO DMAPP TO FORM GERANYLGERANYL PYROPHOSPHATE.
CC   -!- CATALYTIC ACTIVITY: Dimethylallyl diphosphate + isopentenyl
CC       diphosphate = diphosphate + geranyl diphosphate.
CC   -!- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
CC       = diphosphate + trans,trans-farnesyl diphosphate.
CC   -!- CATALYTIC ACTIVITY: Trans-trans-farnesyl diphosphate + isopentenyl
CC       diphosphate = diphosphate + geranylgeranyl diphosphate.
CC   -!- PATHWAY: FIRST COMMITTED STEP IN CAROTENOID BIOSYNTHESIS. KEY
CC       ENZYME IN PLANT TERPENOID BIOSYNTHESIS.
CC   -!- SUBUNIT: MONOMER.
CC   -!- SUBCELLULAR LOCATION: Chloroplast stroma; chromoplast (Probable).
CC   -!- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; X98795; CAA67330.1; -.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
KW   Transferase; Isoprene biosynthesis; Carotenoid biosynthesis;
KW   Chloroplast; Transit peptide.
FT   TRANSIT       1      ?       CHLOROPLAST.
FT   CHAIN         ?    366       GERANYLGERANYL PYROPHOSPHATE SYNTHETASE.
SQ   SEQUENCE   366 AA;  39432 MW;  1EADBFD9C4CF4F77 CRC64;

  Query Match             33.0%;  Score 479;  DB 1;  Length 366;
  Best Local Similarity   42.2%;  Pred. No. 3.6e-28;
  Matches  114;  Conservative   41;  Mismatches  105;  Indels   10;  Gaps    4;

Qy         11 DEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKSAIA 70
              | ||  |  |:       :: |:| ||| |||||:|||| :   : :  |  | | :  |
Db         81 DSVNKALDSAVPLR-EPLKIHEAMRYSLLAGGKRVRPVLCIAACELVGGEESLAMPARCA 139

Qy         71 LEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSDDRLT 130
              :||||| |||||||| |||| |||| |||||||  |:|||||: ||| ::|    :
Db        140 VEMIHTMSLIHDDLPCMDNDDLRRGKPTNHKVYGEDVAVLAGDALLSFAFEHLAS--ATS 197

Qy        131 DEVK----IKVLQRLSIASGHVGMVGGQMLDMQSEG---QPIDLETLEMIHKTKTGALLT 183
              ||      :: :  |: | |  |:| ||::|: |||     : || |: ||  || |||
Db        198 SEVSPARVVRAVGELAKAIGTEGLVAGQVVDISSEGLDLNNVGLEHLKFIHLHKTAALLE 257

Qy        184 FAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKS 243
              : :    |    |   | |  ::  :|::||: ||:||      :|||  | ||  :|
Db        258 ASAVLGGIIGGGSDEEIERLRKFARCIGLLFQVVDDILDVTKSSQELGKTAGKDLIADKL 317

Qy        244 TYVSLLGKDGAEDKLTYHRDAAVDELTQID 273
              ||  |:| : : :         | |:|   |
Db        318 TYPKLMGLEKSREFAEKLNTEARDQLLGFD 347


RESULT 10
GGPP_ARATH
ID   GGPP_ARATH     STANDARD;      PRT;   371 AA.
AC   P34802; O23201;
DT   01-FEB-1994 (Rel. 28, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Geranylgeranyl pyrophosphate synthetase, chloroplast precursor (GGPP
DE   synthetase) (GGPS) [Includes: Dimethylallyltransferase (EC 2.5.1.1);
DE   Geranyltranstransferase (EC 2.5.1.10); Farnesyltranstransferase
DE   (EC 2.5.1.29)].
GN   GGPS1 OR AT4G36810.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94286754; PubMed=8016276;
RA   Scolnik P.A., Bartley G.E.;
RT   "Nucleotide sequence of an Arabidopsis cDNA for geranylgeranyl
RT   pyrophosphate synthase.";
RL   Plant Physiol. 104:1469-1470(1994).
```

```
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083488; PubMed=10617198;
RA   Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA   Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA   Harris B., Ansorge W., Brandt P., Grivell L., Rieger M.,
RA   Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA   Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,
RA   Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA   Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA   Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA   Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA   Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA   Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA   Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA   Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA   Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA   De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
RA   Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA   Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
RA   Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA   Borkova D., Bloecker H., Scharfe M., Grimm M., Loehnert T.-H.,
RA   Dose S., de Haan M., Maarse A., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA   Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
RA   Massenet O., Quigley F., Clabauld G., Muendlein A., Felber R.,
RA   Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA   Chefdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA   Gibbons T., Weber N., Vandenbol M., Bargues M., Terol J., Torres A.,
RA   Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA   Heijnen L., Schwarz S., Scholler P., Heber S., Francs P., Bielke C.,
RA   Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA   Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA   Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA   Sehkon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA   Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA   Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA   Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA   Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA   Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA   Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshu C.,
RA   Antonoiu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA   Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA   Swaby I.K., O'Shaugnessy A., Rodriguez M., Hoffman J., Till S.,
RA   Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA   Chen E., Marra M., Martienssen R., McCombie W.R.;
RT   "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:769-777(1999).
CC   -!- FUNCTION: CATALYZES THE TRANS-ADDITION OF THE THREE MOLECULES OF
CC       IPP ONTO DMAPP TO FORM GERANYLGERANYL PYROPHOSPHATE.
CC   -!- CATALYTIC ACTIVITY: Dimethylallyl diphosphate + isopentenyl
CC       diphosphate = diphosphate + geranyl diphosphate.
CC   -!- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
CC       = diphosphate + trans,trans-farnesyl diphosphate.
CC   -!- CATALYTIC ACTIVITY: Trans-trans-farnesyl diphosphate + isopentenyl
CC       diphosphate = diphosphate + geranylgeranyl diphosphate.
CC   -!- PATHWAY: FIRST COMMITTED STEP IN CAROTENOID BIOSYNTHESIS. KEY
CC       ENZYME IN PLANT TERPENOID BIOSYNTHESIS.
CC   -!- SUBUNIT: MONOMER.
CC   -!- SUBCELLULAR LOCATION: Chloroplast.
CC   -!- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L25813; AAA32797.1; -.
DR   EMBL; Z99708; CAB16803.1; -.
DR   EMBL; AL161590; CAB80347.1; -.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
KW   Transferase; Isoprene biosynthesis; Carotenoid biosynthesis;
KW   Chloroplast; Transit peptide.
FT   TRANSIT       1      ?       CHLOROPLAST.
FT   CHAIN         ?    371       GERANYLGERANYL PYROPHOSPHATE SYNTHETASE.
FT   CONFLICT    108    108       R -> S (IN REF. 1).
FT   CONFLICT    141    141       A -> R (IN REF. 1).
FT   CONFLICT    192    192       A -> S (IN REF. 1).
SQ   SEQUENCE   371 AA;  40174 MW;  EFA8088A75B6A005 CRC64;

  Query Match             32.9%;  Score 478;  DB 1;  Length 371;
  Best Local Similarity   40.5%;  Pred. No. 4.4e-28;
  Matches  111;  Conservative   49;  Mismatches  100;  Indels   14;  Gaps    4;

Qy          6 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGM 65
              :|| :|      :| : : :   :: |:| ||| |||||:|||| :   : :  |    |
Db         87 VNKALDS-----AVPLREPL---KIHEAMRYSLLAGGKRVRPVLCIAACELVGGEESTAM 138

Qy         66 KSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELI-- 123
               :| |:|||| |||||||| ||||| |||| |||||:||  |:|||||||: ||| :
Db        139 PAACAVEMIHTMSLIHDDLPCMDNDDLRRGKPTNHKVFGEDVAVLAGDALLSFAFEHLAS 198

Qy        124 -SSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEG---QPIDLETLEMIHKTKTG 179
               :| | ::    :: :  |: | |  |:| ||::|: |||     : || || ||  ||
Db        199 ATSSDVVSPVRVVVRAVGELAKAIGTEGLVAGQVVDISSEGLDLNDVGLEHLEFIHLHKTA 258

Qy        180 ALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLE 239
              |||  : :  | :    |  | |  ::  :|::||: ||:||      :|||  | ||
Db        259 ALLEASAVLGAIVGGGSDDEIERLRKFARCIGLLFQVVDDILDVTKSSKELGKTAGKDLI 318

Qy        240 NNKSTYVSLLGKDGAEDKLTYHRDAAVDELTQID 273
               :| ||  ::| : : :        | |:|   |
Db        319 ADKLTYPKIMGLEKSREFAEKLNREARDQLLGFD 352


RESULT 11
IDSA_METJA
ID   IDSA_METJA     STANDARD;      PRT;   327 AA.
AC   Q58270;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Bifunctional short chain isoprenyl diphosphate synthase [Includes:
DE   Farnesyl pyrophosphate synthetase (EC 2.5.1.1) (FPP synthetase)
DE   (Dimethylallyltransferase); Geranyltranstransferase (EC 2.5.1.10)].
GN   IDSA OR MJ0860.
OS   Methanococcus jannaschii.
OC   Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC   Methanocaldococcaceae; Methanocaldococcus.
OX   NCBI_TaxID=2190;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX   MEDLINE=96337999; PubMed=8688087;
RA   Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA   Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA   Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA   Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA   Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA   Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA   Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA   Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT   "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT   jannaschii.";
RL   Science 273:1058-1073(1996).
CC   -!- CATALYTIC ACTIVITY: Dimethylallyl diphosphate + isopentenyl
CC       diphosphate = diphosphate + geranyl diphosphate.
CC   -!- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
```

```
CC       = diphosphate + trans,trans-farnesyl diphosphate.
CC    -!- COFACTOR: REQUIRES MAGNESIUM (BY SIMILARITY).
CC    -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC    -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC    -!- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
CC    ---------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    ---------------------------------------------------------------------
DR    EMBL; U67530; AAB98865.1; -.
DR    TIGR; MJ0860; -.
DR    InterPro; IPR000092; Polyprenyl_synt.
DR    Pfam; PF00348; polyprenyl_synt; 1.
DR    PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
DR    PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
KW    Lipid synthesis; Isoprene biosynthesis; Transferase; Magnesium;
KW    Multifunctional enzyme; Complete proteome.
SQ    SEQUENCE   327 AA;  37568 MW;  07D68AC9BD657DAC CRC64;

  Query Match             30.1%;  Score 438;  DB 1;  Length 327;
  Best Local Similarity   41.2%;  Pred. No. 3.2e-25;
  Matches  115;  Conservative   45;  Mismatches   87;  Indels   32;  Gaps    9;

Qy          9 LIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKSA 68
              :: ::: ||   ::|   | :|  :  : | |||||||| | ::|      : |  : :|
Db         17 ILQKIDEELKTYVDK---DDKLYNASKHLLFAGGKRIRPYLTVVTYMLKKDDIEEVLPAA 73

Qy         69 IALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSDDR 128
               |:|:|| |:|||||:  ||||| |||| | | ||||  |||||| |  |||| :|   |
Db         74 AAVELIHNYTLIHDDI--MDNDDERRGKPTVHVVYGEPMAILAGDLLYAKAFEAVS---R 128

Qy        129 LTDEVKI-KVLQRLSIASGHVGMVGGQMLDMQSEG-QPIDLETLEMIHKTKTGALLTFAV 186
              : |  |  :||: || |   | :  || :||: |   |   | |:|| | ||||||  :|
Db        129 IKDNKKAHEVLKILSKAC--VEVCEGQAMDMEFENYYPTMEEYLDMIRK-KTGALLEASV 185

Qy        187 MSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYV 246
                | :|: ::  :| |: |:  :|: |||:||:||  ||: |||| ||||:   | | :
Db        186 GIGAVMADCNEEEREALKEYAKRIGLTFQIQDDVLDLIGDQKKLGKPVGSDIREGKKTII 245

Qy        247 SLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIV 285
                          :  |  :||    | ||||:
Db        246 -----------------VIHALKTLDED-KKKRLLEIL 265
```

```
ULT 12
CRTE_PANAN
ID   CRTE_PANAN     STANDARD;      PRT;   302 AA.
AC   P21684;
DT   01-MAY-1991 (Rel. 18, Created)
DT   01-MAY-1991 (Rel. 18, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Geranylgeranyl pyrophosphate synthetase (EC 2.5.1.29) (GGPP
DE   synthetase) (Farnesyltranstransferase).
GN   CRTE.
OS   Pantoea ananas (Erwinia uredovora).
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Pantoea.
OX   NCBI_TaxID=553;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=20D3;
RX   MEDLINE=91072214; PubMed=2254247;
RA   Misawa N., Nakagawa M., Kobayashi K., Yamano S., Izawa Y.,
RA   Nakamura K., Harashima K.;
RT   "Elucidation of the Erwinia uredovora carotenoid biosynthetic pathway
RT   by functional analysis of gene products expressed in Escherichia
RT   coli.";
RL   J. Bacteriol. 172:6704-6712(1990).
```

```
CC    -!- CATALYTIC ACTIVITY: Trans-trans-farnesyl diphosphate + isopentenyl
CC        diphosphate = diphosphate + geranylgeranyl diphosphate.
CC    -!- PATHWAY: Carotenoid biosynthesis.
CC    -!- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
CC    ---------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    ---------------------------------------------------------------------
DR    EMBL; D90087; BAA14124.1; -.
DR    PIR; A37802; A37802.
DR    InterPro; IPR000092; Polyprenyl_synt.
DR    Pfam; PF00348; polyprenyl_synt; 1.
DR    PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
DR    PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
KW    Carotenoid biosynthesis; Isoprene biosynthesis; Transferase.
SQ    SEQUENCE   302 AA;  32583 MW;  CADB04699D2EBA4A CRC64;

  Query Match             25.9%;  Score 376.5;  DB 1;  Length 302;
  Best Local Similarity   35.6%;  Pred. No. 9.5e-21;
  Matches  100;  Conservative   48;  Mismatches  118;  Indels   15;  Gaps    5;

Qy          6 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLN--TEYEL 63
              ::: :|::   | |   : |:   : | |    | ||||||:|||||   |     ::
Db         23 IDRRLDQL---LPVEGERDVVGAAMREGAL----APGKRIRPMLLLLTARDLGCAVSHDG 75

Qy         64 GMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELI 123
               :  | |:||:|  ||| ||:| ||:   |||: | |  ||| ||||  |||:||| :|
Db         76 LLDLACAVEMVHAASLILDDMPCMDDAKLRRGRPTIHSHYGEHVAILAAVALLSKAFGVI 135

Qy        124 SSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLT 183
              :  | ||   | : :  || | |  |:| ||  |:    :|   | : | :  || |
Db        136 ADADGLTPLAKNRAVSELSNAIGMQGLVQGQFKDLSEGDKPRSAEAILMTNHFKTSTLFC 195

Qy        184 FAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKS 243
              ::  |: :||     :: |  :|  ||  ||: ||| |   |  |      |: :  ||
Db        196 ASMQMASIVANASSEARDCLHRFSLDLGQAFQLLDDLTDGMTDTGK-----DSNQDAGKS 250

Qy        244 TYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFN-TKHLLE 283
              | |:||    |::|  |   | : |:   :  : |:| ::
Db        251 TLVNLLGPRAVEERLRQHLQLASEHLSAACQHGHATQHFIQ 291
```

```
RESULT 13
IDSA_METTM
ID   IDSA_METTM     STANDARD;      PRT;   324 AA.
AC   Q53479;
DT   01-NOV-1997 (Rel. 35, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Bifunctional short chain isoprenyl diphosphate synthase [Includes:
DE   Farnesyl pyrophosphate synthetase (EC 2.5.1.1) (FPP synthetase)
DE   (Dimethylallyltransferase); Geranyltranstransferase (EC 2.5.1.10)].
GN   IDSA.
OS   Methanobacterium thermoautotrophicum (strain Marburg / DSM 2133).
OC   Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC   Methanobacteriaceae; Methanothermobacter.
OX   NCBI_TaxID=79929;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=95070145; PubMed=7979381;
RA   Chen A., Poulter C.D.;
RT   "Isolation and characterization of idsA: the gene for the short chain
RT   isoprenyl diphosphate synthase from Methanobacterium
RT   thermoautotrophicum.";
RL   Arch. Biochem. Biophys. 314:399-404(1994).
CC   -!- CATALYTIC ACTIVITY: Dimethylallyl diphosphate + isopentenyl
CC       diphosphate = diphosphate + geranyl diphosphate.
CC   -!- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
```

```
CC       = diphosphate + trans,trans-farnesyl diphosphate.
CC   -!- COFACTOR: REQUIRES MAGNESIUM.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S75695; AAB32421.1; -.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
KW   Lipid synthesis; Isoprene biosynthesis; Transferase; Magnesium;
KW   Multifunctional enzyme.
FT   INIT_MET      0      0
SQ   SEQUENCE   324 AA;  35505 MW;  88558AE4DF5F66C1 CRC64;

  Query Match             25.9%;  Score 376.5;  DB 1;  Length 324;
  Best Local Similarity   35.4%;  Pred. No. 1e-20;
  Matches  104;  Conservative   53;  Mismatches  110;  Indels   27;  Gaps   10;

Qy          6 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGM 65
              : |   :   :   :    |:    || |: |   : : ||||:||| | ||: :::    |
Db          7 LRKYSEVADKRIMECISDITPDTLLKASE-HLITAGGKKIRPSLALLSCEAVGGNPEDAA 65

Qy         66 KSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELI-- 123
                | |:|:|||:|||||:  ||:|: |||: : | ::||  |||||| | :|||| :
Db         66 GVAAAIELIHTFSLIHDDI--MDDDEMRRGEPSVHVIWGEPMAILAGDVLFSKAFEAVIR 123

Qy        124 -SSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDL---ETLEMIHKTKTG 179
               :|: | : :       :    | :  || |||  | : :|:   | :|||:| ||
Db        124 NGDSERVKDALAV-------VVDSCVKICEGQALDMGFE-ERLDVTEDEYMEMIYK-KTA 174

Qy        180 ALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLE 239
              ||:  |  : | :   :   ||||  |    :|: ||| || ||   ||  ||| ||||:
Db        175 ALIAAATKAGAIMGGASEREVEALEDYGKFIGLAFQIHDDYLDVVSDEESLGKPVGSDIA 234

Qy        240 NNKST--YVSLLGKDGAEDK------LTYHRDAAVDELTQIDEQFN-TKHLLEI 284
                | |   |  | :   ||:      |    : :| |  :| |::  |::  |:
Db        235 EGKMTLMVVKALEEASEEDRERLISILGSGDEGSVAEAIEIFERYGATQYAHEV 288


RESULT 14
ISPA_BUCAI
ID   ISPA_BUCAI     STANDARD;      PRT;   282 AA.
AC   P57537;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Geranyltranstransferase (EC 2.5.1.10) (Farnesyl-diphosphate synthase)
DE   (FPP synthase).
GN   ISPA OR BU465.
OS   Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS   symbiotic bacterium).
OC   Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX   NCBI_TaxID=118099;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Tokyo 1998;
RX   MEDLINE=20445173; PubMed=10993077;
RA   Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT   "Genome sequence of the endocellular bacterial symbiont of aphids
RT   Buchnera sp. APS.";
RL   Nature 407:81-86(2000).
CC   -!- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
CC       = diphosphate + trans,trans-farnesyl diphosphate.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AP001119; BAB13162.1; -.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
KW   Transferase; Isoprene biosynthesis; Complete proteome.
SQ   SEQUENCE   282 AA;  32331 MW;  228B31A6DBCD6445 CRC64;

  Query Match             25.8%;  Score 375;  DB 1;  Length 282;
  Best Local Similarity   35.6%;  Pred. No. 1.1e-20;
  Matches  93;  Conservative   55;  Mismatches   83;  Indels   30;  Gaps   7;

Qy         13 VNNELSVAINK-SVMDTQLEESMLYSLNAGGKRIRPVLL----------LLTLDSLNTEY 61
              :| :|   :|:     : | ::| ||: :| ||:|  |:          : ||| ::|
Db         14 INQKLFYTLNQLPFQKSSLLKAMKYSVFSGSKRLRSSLIYSTGDVFKVNITTLDVIST-- 71

Qy         62 ELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFE 121
                    |:| ||:|||||||| ||||::||||:: |  ||| |::|||||| : ||
Db         72 --------AIEFIHSYSLIHDDLPCMDNDNFRRGKISCHVKYGESTSLLAGDALQSLAFN 123

Qy        122 LISSD--DRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTG 179
              ::|:     :::  :||::  || : |  ||  || ||:::| : ::|  ||:|:  ||
Db        124 ILSNSFMPNVSNLKRIKMISELSYSIGSSGMCMGQNLDLEAEKKDVNLSELEIINLYKTS 183

Qy        180 ALLTFAV-MSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKL-GKKVGSD 237
              |:  || :       |   :    |: :|  :|: |||:||:||   |  |   ||:
Db        184 FLMRSAVRLVYFSSNNFSKSILSILDLFSISIGLAFQIQDDILDFKKDSVKTDNKKI--- 240

Qy        238 LENNKSTYVSLLGKDGAEDKL 258
                | ||  ::| | :  |:
Db        241 --IKKHTYPLIIGLDESRKKI 259


RESULT 15
ISPA_RHISN
ID   ISPA_RHISN     STANDARD;      PRT;   332 AA.
AC   P55539;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Probable geranyltranstransferase (EC 2.5.1.10) (Farnesyl-diphosphate
DE   synthase) (FPP synthase).
GN   Y4KU.
OS   Rhizobium sp. (strain NGR234).
OG   Plasmid sym pNGR234a.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Rhizobium.
OX   NCBI_TaxID=394;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97305956; PubMed=9163424;
RA   Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA   Perret X.;
RT   "Molecular basis of symbiosis between Rhizobium and legumes.";
RL   Nature 387:394-401(1997).
CC   -!- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
CC       = diphosphate + trans,trans-farnesyl diphosphate.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AE000082; AAB91752.1; -.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
KW   Transferase; Isoprene biosynthesis; Plasmid.
SQ   SEQUENCE   332 AA;  34688 MW;  EE68C1547449AA87 CRC64;

  Query Match             24.2%;  Score 352;  DB 1;  Length 332;
  Best Local Similarity   36.5%;  Pred. No. 6.7e-19;
  Matches   92;  Conservative   40;  Mismatches  108;  Indels   12;  Gaps    3;

Qy         27 DTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKSAIALEMIHTYSLIHDDLPA 86
              :|:|  :| |:   |||| | :| |      :|   :       |:||:|  :|:||||||
Db         59 ETELMAAMRYATLHGGKRTRALLCLAAGALADTPAHILDDVGAAIEMMHACTLVHDDLPA 118

Qy         87 MDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSDDRLTDEVKIKVLQRLSIASG 146
              ||:|  |||  | |  :|| |||| ||||   ||  ::| |   |  :| ::: |: |
Db        119 MDDDVLRRGLPTVHVKFGEATAILVGDALQAHAFLTLASLDAPGDN-RIALVRELAQAVS 177

Qy        147 HVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFAVMSAADIANVDD----TTKEH 202
             |  ||| :|:   |: ::|: :  :|: | |||:  :|   |  |  :|    |
Db        178 AEGAAGGQAMDLSLVGKHVELDRIVAMHRMKCGALVRASVRMGALCAIAEDAADATLYCA 237

Qy        203 LESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVSLLGKDGAEDKLTYHR 262
             |: ||   |:  |: ||:||   | | |||  | |    | |  |::|   |       |
Db        238 LDHYSACFGLALQVVDDILDATADTATLGKTPGKDAAAQKPTCASIMGLQAA-------R 290

Qy        263 DAAVDELTQIDE 274
              |:| | :  |
Db        291 QFALDLLCEAGE 302


Search completed: May 29, 2003, 11:59:11
Job time : 12 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 29, 2003, 11:54:20 ; Search time 33 Seconds
(without alignments)
1791.986 Million cell updates/sec

Title: US-09-925-637-64
Perfect score: 1453
Sequence: 1 MTNLPMNKLIDEVNNELSVA.....ELTQIDEQFNTKHLLEIVDL 287

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1444	99.4	293	16	Q99TX2	Q99tx2 staphyloco
2	702.5	48.3	293	16	Q8Y7C2	Q8y7c2 listeria mo
3	686.5	47.2	293	16	Q92BZ1	Q92bz1 listeria in
4	638	43.9	294	16	Q9K969	Q9k969 bacillus ha
5	576	39.6	297	16	Q8RE07	Q8re07 fusobacteri
6	551.5	38.0	290	16	Q99YX5	Q99yx5 streptococc
7	551	37.9	301	16	Q8XJE0	Q8xje0 clostridium
8	549	37.8	291	16	Q97QK0	Q97qk0 streptococc
9	540.5	37.2	285	16	Q9CH81	Q9ch81 lactococcus
10	535.5	36.9	383	10	Q8W1R9	Q8w1r9 abies grand
11	534.5	36.8	309	16	Q8Z085	Q8z085 anabaena sp
12	534	36.8	307	2	Q9S5F1	Q9s5f1 synechococc
13	532	36.6	289	16	Q97HD2	Q97hd2 clostridium
14	526.5	36.2	302	16	P72683	P72683 synechocyst
15	520.5	35.8	393	10	Q9ZPM3	Q9zpm3 taxus canad
16	519	35.7	367	10	Q9SXZ6	Q9sxz6 daucus caro
17	516.5	35.5	294	16	Q8RAC7	Q8rac7 thermoanaer
18	516.5	35.5	366	10	Q9SXZ5	Q9sxz5 daucus caro
19	499.5	34.4	356	10	O81099	O81099 helianthus
20	489	33.7	361	10	Q9FV47	Q9fv47 tagetes ere
21	485.5	33.4	298	16	Q9JSM0	Q9jsm0 neisseria m
22	485.5	33.4	299	16	Q8XE75	Q8xe75 escherichia
23	484	33.3	294	16	Q9KTL2	Q9ktl2 vibrio chol
24	483.5	33.3	288	10	Q94IF0	Q94if0 eucommia ul
25	482.5	33.2	370	10	Q94ID7	Q94id7 hevea brasi
26	481	33.1	295	16	Q9CNA1	Q9cna1 pasteurella
27	480.5	33.1	298	2	Q9RME5	Q9rme5 zymomonas m
28	478.5	32.9	316	10	Q42866	Q42866 lupinus alb
29	478.5	32.9	368	10	Q9SSU0	Q9ssu0 croton subl
30	475.5	32.7	306	16	Q8ZC46	Q8zc46 yersinia pe
31	474.5	32.7	295	16	Q9HWY4	Q9hwy4 pseudomonas
32	473	32.6	299	16	Q8ZRD0	Q8zrd0 salmonella
33	473	32.6	299	16	Q8Z8X2	Q8z8x2 salmonella
34	473	32.6	360	10	Q9LUE1	Q9luel arabidopsis
35	470.5	32.4	294	16	Q9A6M4	Q9a6m4 caulobacter
36	469.5	32.3	259	16	Q9K1A6	Q9k1a6 neisseria m
37	469	32.3	376	10	Q9XHX1	Q9xhx1 oryza sativ
38	467	32.1	347	10	Q9ZU77	Q9zu77 arabidopsis
39	465.5	32.0	377	10	Q9SBR3	Q9sbr3 mentha pipe
40	462	31.8	299	16	Q8XX94	Q8xx94 ralstonia s
41	462	31.8	360	10	Q9LUD9	Q9lud9 arabidopsis
42	457.5	31.5	335	16	Q8UBX7	Q8ubx7 agrobacteri
43	456	31.4	360	10	Q9LHR4	Q9lhr4 arabidopsis
44	452	31.1	357	10	Q9LIA0	Q9lia0 arabidopsis
45	450.5	31.0	306	10	Q9LRR0	Q9lrr0 arabidopsis

ALIGNMENTS

RESULT 1
Q99TX2
ID Q99TX2 PRELIMINARY; PRT; 293 AA.
AC Q99TX2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE IspA protein (Geranyltranstransferase homolog).
GN ISPA OR SAV1521 OR SA1352.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of meticillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003362; BAB57683.1; -.
DR EMBL; AP003134; BAB42614.1; -.
DR InterPro; IPR000092; Polyprenyl_synt.
DR Pfam; PF00348; polyprenyl_synt; 1.
DR PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
KW Transferase; Complete proteome.
SQ SEQUENCE 293 AA; 32746 MW; E7AFFBBBA868F7D8 CRC64;

Query Match 99.4%; Score 1444; DB 16; Length 293;
Best Local Similarity 99.3%; Pred. No. 8.6e-103;

```
  Matches  285;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MTNLPMNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTNLPMNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTE 60

Qy         61 YELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAF 120
              ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YELGVKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAF 120

Qy        121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGA 180

Qy        181 LLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLEN 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        181 LLTFAVMSAADIANVDDATKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLEN 240

Qy        241 NKSTYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NKSTYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287


RESULT 2
Q8Y7C2
ID   Q8Y7C2      PRELIMINARY;      PRT;   293 AA.
AC   Q8Y7C2;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical protein lmo1363.
GN   LMO1363.
OS   Listeria monocytogenes.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Listeriaceae; Listeria.
OX   NCBI_TaxID=1639;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EGD-E / SEROVAR 1/2A;
RX   MEDLINE=21537279; PubMed=11679669;
RA   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA   Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA   Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA   Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA   Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA   Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA   Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA   Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA   Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT   "Comparative genomics of Listeria species.";
RL   Science 294:849-852(2001).
DR   EMBL; AL591978; CAC99441.1; -.
DR   ListiList; LMO01363; -.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   InterPro; IPR000235; Ribosomal_S7.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
DR   PROSITE; PS00052; RIBOSOMAL_S7; UNKNOWN_1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   293 AA;  32308 MW;  6A7F0C75B1824CC9 CRC64;

  Query Match             48.3%;  Score 702.5;  DB 16;  Length 293;
  Best Local Similarity   52.9%;  Pred. No. 5.6e-46;
  Matches  148;  Conservative   46;  Mismatches   81;  Indels    5;  Gaps    2;

Qy          8 KLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKS 67
              |:|||    |   ||:  :: :|:|||||||: |||||||||:|:  || :|     ||:|:
Db         13 KVIDE---SLFKEINERNIEPRLKESMLYSIQAGGKRIRPMLVFATLQALKVNPLLGVKT 69

Qy         68 AIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSDD 127
              | ||||||||||||||||||||||||||||| |||||:|: |||||||||| || ::: ||
Db         70 ATALEMIHTYSLIHDDLPAMDNDDYRRGKWTNHKVFGDATAILAGDALLTLAFSILAEDD 129

Qy        128 RLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFAVM 187
              |: | :| :: ::| :|| ||||||: |:::| : : || |  ||  ||| || :||
Db        130 NLSFETRIALINQISFSSGAEGMVGGQLADLEAENKQVTLEELSSIHARKTGELLIYAVT 189

Qy        188 SAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVS 247
              ||| ||  |   : |  :: ::|: ||| ||:||  ||| |:||| |:|   |||||
Db        190 SAAKIAEADPEQTKRLRIFAENIGIGFQISDDILDVIGDETKMGKKTGADAFLNKSTYPG 249

Qy        248 LLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287
              ||  |||:  |  |   |  |:  |  |: : ||:: ||
Db        250 LLTLDGAKRALNEHVTIAKSALSGHD--FDDEILLKLADL 287


RESULT 3
Q92BZ1
ID   Q92BZ1      PRELIMINARY;      PRT;   293 AA.
AC   Q92BZ1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Hypothetical protein lin1400.
GN   LIN1400.
OS   Listeria innocua.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Listeriaceae; Listeria.
OX   NCBI_TaxID=1642;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CLIP 11262 / SEROVAR 6A;
RX   PubMed=11679669;
RA   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA   Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA   Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA   Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA   Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA   Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA   Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA   Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA   Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT   "Comparative genomics of Listeria species.";
RL   Science 294:849-852(2001).
DR   EMBL; AL596168; CAC96631.1; -.
DR   ListiList; LIN01400; -.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   InterPro; IPR000235; Ribosomal_S7.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; UNKNOWN_1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; UNKNOWN_1.
DR   PROSITE; PS00052; RIBOSOMAL_S7; UNKNOWN_1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   293 AA;  32459 MW;  36649E7E47D8988D CRC64;

  Query Match             47.2%;  Score 686.5;  DB 16;  Length 293;
  Best Local Similarity   52.1%;  Pred. No. 9.5e-45;
  Matches  146;  Conservative   45;  Mismatches   84;  Indels    5;  Gaps    2;

Qy          8 KLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKS 67
              |::||    |   ||   :: :|:|||||: ||||||||:|:  || :|| |   |:|:
Db         13 KVLDE---SLFKEINMRNIEPKLKESMLYSVQAGGKRIRPMLVFATLQALNIEPMRGLKT 69

Qy         68 AIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSDD 127
              | |||||||||||||||||||||:||||||| ||||:||: |||||||||| || ::: |:
Db         70 ATALEMIHTYSLIHDDLPAMDDDDYRRGKWTNHKIYGDATAILAGDALLTLAFSILAEDE 129

Qy        128 RLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFAVM 187
              |: | :| :: ::| :||  ||||||   ||::| : : || |  ||  ||| || |||
Db        130 NLSFETRIALINQISYSSGAEGMVGGQQADMEAENKQVTLEELASIHARKTGELLIFAVT 189
```

```
Qy        188 SAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVS 247
              |||  ||       : |  :: ::|: ||| ||:||  ||| |:||| | |   |||||
Db        190 SAAKIAEATPEQTKRLRIFAENIGIGFQISDDILDVIGDETKMGKKTGVDAFLNKSTYPG 249

Qy        248 LLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287
              ||  :||:  |  |   |   |:  |  |: : ||:: ||
Db        250 LLTLEGAKRALNEHVSIAKSALSGHD--FDDEILLKLADL 287


RESULT 4
Q9K969
ID   Q9K969      PRELIMINARY;      PRT;   294 AA.
AC   Q9K969;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Geranyltranstransferase.
GN   BH2781.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Bacillaceae; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
DR   EMBL; AP001516; BAB06500.1; -.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
KW   Transferase; Complete proteome.
SQ   SEQUENCE   294 AA;  32653 MW;  3B2942868A62AD9E CRC64;

  Query Match           43.9%;  Score 638;  DB 16;  Length 294;
  Best Local Similarity  47.7%;  Pred. No. 5e-41;
  Matches  134;  Conservative   53;  Mismatches   92;  Indels    2;  Gaps    1;

Qy          6 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGM 65
              :::: | :   :   | :      |: |||||| |||||||| ||| |: |   :   |:
Db          9 LDEIKDIIEERMPAHIERLNSPDMLKNSMLYSLKAGGKRIRPALLLATMKSFQKDISQGI 68

Qy         66 KSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISS 125
               | |:||||||||||||||:||:|| |||| |||||:||  |||||||||| :||:::
Db         69 DLACAIEMIHTYSLIHDDLPSMDDDDIRRGKPTNHKVFGEAHAILAGDALLTYSFEIVAK 128

Qy        126 DDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFA 185
              :   :  : ::: |: |:|  |||||: |:: | | : :| || ||  ||||||:||
Db        129 MKGVDPAKTLCLIEELARAAGPEGMVGGQVADIEGENQKLTVEGLEYIHHHKTGALLSFA 188

Qy        186 VMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTY 245
              ::: | :|:  :   |::  :|  ||::|||||:||  ||:| :|| |||  | ||||
Db        189 IVAGARLADASEQDIENIRRFSRELGLLFQIKDDILDVEGDQAAIGKPVGSDDGNQKSTY 248

Qy        246 VSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVD 286
              |||  :||::||  |   | :.|  :  | | : | |: |
Db        249 PSLLTLEGAKEKLHLHTLLAKEYLYSV--QMNHRLLEELTD 287


RESULT 5
Q8RE07
ID   Q8RE07      PRELIMINARY;      PRT;   297 AA.
AC   Q8RE07;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Dimethylallyltransferase (EC 2.5.1.1).
GN   FN1327.
OS   Fusobacterium nucleatum (subsp. nucleatum).
OC   Bacteria; Fusobacteria; Fusobacterium.
OX   NCBI_TaxID=76856;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 25586;
RX   MEDLINE=21886394; PubMed=11889109;
RA   Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA   Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA   Vasieva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA   Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA   Fonstein M., Kyrpides N., Overbeek R.;
RT   "Genome sequence and analysis of the oral bacterium Fusobacterium
RT   nucleatum strain ATCC 25586.";
RL   J. Bacteriol. 184:2005-2018(2002).
DR   EMBL; AE010637; AAL95523.1; -.
KW   Transferase; Complete proteome.
SQ   SEQUENCE   297 AA;  33577 MW;  89DC35727945FC53 CRC64;

  Query Match           39.6%;  Score 576;  DB 16;  Length 297;
  Best Local Similarity  44.2%;  Pred. No. 2.8e-36;
  Matches  123;  Conservative   54;  Mismatches   99;  Indels    2;  Gaps    1;

Qy         11 DEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKSAIA 70
              |    ||   : :      : : | |::  ||||:|| || :||: ||     |:|||||
Db         13 DFFETELKKELKELSYPETIAKGMEYAILNGGKRLRPFLLFVTLELLNENINKGVKSAIA 72

Qy         71 LEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSD--DR 128
              |||||:|||:||||||:|||||||||||| |||:||   || ||:||| || ::|    :
Db         73 LEMIHSYSLVHDDLPALDNDDYRRGKLTTHKVFGEAEGILIGDSLLTYAFYVLSQKNLEL 132

Qy        129 LTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFAVMS 188
              |: :  : :: : |  :|  ||:||||:|:||| : |||||: ||  ||| |:   :
Db        133 LSSKQIVNIISKTSEYAGIDGMIGGQMIDIQSENKKIDLETLKYIHSHKTGKLIKLPIEI 192

Qy        189 AADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVSL 248
              |  |||::   :| || |:  :|: ||:|||:|| |    ||| |||:: :|:|| |:
Db        193 ACIIANLEKDKREVLEEYADLIGLAFQVKDDILDVEGTFEDLGKPVGSDVDLHKATYPSI 252

Qy        249 LGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVD 286
              || : ::  |    : | : :     :   | |: : |
Db        253 LGMEESKKILNNTVEKAKEIIKNKFGEEKGKVLISLAD 290


RESULT 6
Q99YX5
ID   Q99YX5      PRELIMINARY;      PRT;   290 AA.
AC   Q99YX5;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Putative geranyltranstransferase (farnesyl diphosphate synthase)
DE   (EC 2.5.1.10).
GN   FPS OR SPY1498.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC   Streptococcaceae; Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;
RX   MEDLINE=21192684; PubMed=11296296;
RA   Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA   Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA   Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J.,
RA   Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT   "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR   EMBL; AE006583; AAK34297.1; -.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
```

```
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
KW   Transferase; Complete proteome.
SQ   SEQUENCE   290 AA;  31824 MW;  875FA2DAE53A0331 CRC64;

  Query Match             38.0%;  Score 551.5;  DB 16;  Length 290;
  Best Local Similarity   44.8%;  Pred. No. 2.1e-34;
  Matches  130;  Conservative   46;  Mismatches   99;  Indels   15;  Gaps    6;

Qy          6 MNKL--IDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSL-----N 58
              |:||  |||          :  | : :| :::|||:::|||||||::||  ::       |
Db          1 MDKLARIDEAIRRYYKTTSNGVSE-ELIDAILYSVDSGGKRIRPLILLEMIEGFGVSLQN 59

Qy         59 TEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTK 118
               ::|    | |||||||| ||||||||||||||||||:||||| :|| ||||||||:|
Db         60 AHFDL----AAALEMIHTGSLIHDDLPAMDNDDYRRGRLTNHKQFGEATAILAGDSLFLD 115

Qy        119 AFELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKT 178
              | ||:  : |  |||: ::| ||:|||  ||||||||||: | | : |  | :||  ||
Db        116 PFGLIAQAE-LNSEVKVALIQELSLASGTFGMVGGQMLDMKGENQALSLPQLSLIHLNKT 174

          179 GALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDL 238
              | || |   :|| |      | :: ||      :|  |||:||:|| |      |||    ||
Db        175 GKLLAFPFKAAALITEQAMTVRQQLEQAGMLIGHAFQIRDDILDVTASFEDLGKTPKKDL 234

Qy        239 ENNKSTYVSLLGKDGAEDKLTYHRDAA--VDELTQIDEQFNTKHLLEIVD 286
                :|| |||| : :   ||   | |  : :  : |  |  : : ::::
Db        235 FAEKATYPSLLGLEASYQLLTESLDQALTIFQTLESDVGFKPQIITKLIE 284


RESULT 7
Q8XJE0
ID   Q8XJE0      PRELIMINARY;      PRT;   301 AA.
AC   Q8XJE0;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Geranyltranstransferase.
GN   CPE1820.
OS   Clostridium perfringens.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1502;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=13 / TYPE A;
RX   PubMed=11792842;
RA   Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA   Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
     "Complete genome sequence of Clostridium perfringens, an anaerobic
     flesh-eater.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR   EMBL; AP003191; BAB81526.1; -.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
KW   Transferase; Complete proteome.
SQ   SEQUENCE   301 AA;  33760 MW;  3BF3494AEF836461 CRC64;

  Query Match             37.9%;  Score 551;  DB 16;  Length 301;
  Best Local Similarity   41.8%;  Pred. No. 2.4e-34;
  Matches  118;  Conservative   60;  Mismatches  100;  Indels    4;  Gaps    3;

Qy          2 TNLPMNKLIDEVNNELSVAINKS-VMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTE 60
              | : :| | :||:  |    ||    :  | :|| ||:| |||||||:|:||:     ::
Db         10 TAMNINSLKEEVDQSLKAYFNKDREYNKVLYDSMAYSINVGGKRIRPILMLLSYYIYKSD 69

Qy         61 YELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAF 120
              |:  :  |:|:|||||||||||||| ||||| |||| |||||:||  |:|||||| :|
Db         70 YKKILTPAMAIEMIHTYSLIHDDLPCMDNDDLRRGKPTNHKVFGEAIAVLAGDALLNEAM 129

Qy        121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEG-QPIDLETLEMIHKTKTG 179
              :::        :  :|  : :: |:|  ||:|||::|: :|  : | |: |: :|  |||

Db        130 KILVDYSLEEGKSALKATKIIADAAGSDGMIGGQIVDIINEDKEEISLKELDYMHLKKTG 189

Qy        180 ALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLE 239
               |:  ::|| | :|   :   : || : | ||: ||||||:||  |:   ||| |  | |
Db        190 ELIKASIMSGAVLAEASEGDIKKLEGFGYKLGLAFQIKDDILDVVGNAKDLGKNVHKDQE 249

Qy        240 NNKSTYVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHL 281
              :||: |::: | :  : |      :  :: |: |  : ||: |
Db        250 SNKNNYIITIFGLEECKKKCVNITEECIEILSSI--KGNTEPL 289


RESULT 8
Q97QK0
ID   Q97QK0      PRELIMINARY;      PRT;   291 AA.
AC   Q97QK0;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Geranyltranstransferase.
GN   SP1205.
OS   Streptococcus pneumoniae.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC   Streptococcaceae; Streptococcus.
OX   NCBI_TaxID=1313;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TIGR4;
RX   MEDLINE=21357209; PubMed=11463916;
RA   Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA   Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA   Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA   Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA   Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA   McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA   Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA   Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT   "Complete genome sequence of a virulent isolate of Streptococcus
RT   pneumoniae.";
RL   Science 293:498-506(2001).
DR   EMBL; AE007421; AAK75312.1; -.
DR   TIGR; SP1205; -.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
KW   Transferase; Complete proteome.
SQ   SEQUENCE   291 AA;  31868 MW;  CF372A7642934062 CRC64;

  Query Match             37.8%;  Score 549;  DB 16;  Length 291;
  Best Local Similarity   46.2%;  Pred. No. 3.2e-34;
  Matches  121;  Conservative   50;  Mismatches   87;  Indels    4;  Gaps    3;

Qy         28 TQLEESMLYSLNAGGKRIRPVLLLLTLDSLN-TEYELGMKSAIALEMIHTYSLIHDDLPA 86
              : | ||:|||::|||||||| |||  |::|  |      : | ||||||| ||||||||
Db         25 SSLRESVLYSIHAGGKRIRPFLLLEVLEALQVTIKPAHAQVATALEMIHTGSLIHDDLPA 84

Qy         87 MDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELISSDDRLTDEVKIKVLQRLSIASG 146
              ||:||||||:|||| :||  ||||||||   :: ||:  | |  ::|: ::  ||:|||
Db         85 MDDDDYRRGRLTNHKKFGEAMAILAGDALFLDSYALIAQAD-LPSQIKVDLIANLSLASG 143

Qy        147 HVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTFAVMSAADIANVDDTTKEHLESY 206
              :||| ||:|||: | | : || |: ||  ||| || :   :|| || :    :  |::
Db        144 SLGMVAGQVLDMEGEHQHLSLEELQTIHANKTGKLLAYPFQAAAIIAELSPEMQVKLKTV 203

Qy        207 SYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVSLLGKDGAEDKLTYHRDAAV 266
              :|: ||::||:||      ::||    ||:  |||| :||| : :         | |
Db        204 GELIGLAFQVRDDVLDVTASFEEIGKTPQKDLQAEKSTYPALLGLEESIAFCNQTLDQAN 263

Qy        267 DELTQIDEQ--FNTKHLLEIVD 286
              ::| :| :|  | |: :: :|:
Db        264 EKLEEIAQQLPFETESIVSVVE 285
```

```
RESULT 9
Q9CH81
ID   Q9CH81      PRELIMINARY;      PRT;   285 AA.
AC   Q9CH81;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Farnesyl diphosphate synthase (EC 2.5.1.10).
GN   ISPA OR LL0857.
OS   Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC   Streptococcaceae; Lactococcus.
OX   NCBI_TaxID=1360;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IL1403;
RX   MEDLINE=21235186; PubMed=11337471;
RA   Bolotin A., Wincker P., Mauger S., Jaillon O., Malarme K.,
RA   Weissenbach J., Ehrlich S.D., Sorokin A.;
RT   "The complete genome sequence of the lactic acid bacterium Lactococcus
RT   lactis ssp. lactis IL1403.";
RT   Genome Res. 11:731-753(2001).
DR   EMBL; AE006320; AAK04955.1; -.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
KW   Complete proteome.
SQ   SEQUENCE   285 AA;  31354 MW;  8678F7A51B850C1F CRC64;

  Query Match             37.2%;  Score 540.5;  DB 16;  Length 285;
  Best Local Similarity   44.8%;  Pred. No. 1.4e-33;
  Matches  130;  Conservative   42;  Mismatches   97;  Indels   21;  Gaps    5;

Qy          6 MNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTE----- 60
              |:  | :: : |:     :   | | ||  ||| ||||||||:| |  |:: : |
Db          1 MDTKILKLEDFLTEFYESAEFPTGLAESAKYSLLAGGKRIRPLLFLNLLEAFDLELSKAH 60

Qy         61 YELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAF 120
              | :    | ||||||| |||||||||||||||||||:||||| : | ||||||| |    |
Db         61 YHV----AAALEMIHTGSLIHDDLPAMDNDDYRRGQLTNHKKFDEATAILAGDTLFFDPF 116

Qy        121 ELISSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGA 180
               ::|: | |: |: : : : |:|||  ||| ||:|||  ||: : |  :| ||: |||
Db        117 FILSTAD-LSAEIIVALTRELAFASGSYGMVAGQILDMAGEGKELTLAEIEQIHRLKTGR 175

Qy        181 LLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLEN 240
              ||||  ::|  :|      | |      ||: |||:||:||     |:||| | |:
Db        176 LLTFPFVAAGIVAQKSTDEVEKLRQVGQILGLAFQIRDDILDVTATFAELGKTPGKDILE 235

          241 NKSTYVSLLGKDGAEDKLTYH--------RDAAV---DELTQIDEQFNTK 279
              |||||: || :||:  || :         | :|    |: :| ||    |
Db        236 EKSTYVAHLGLEGAKKSLTGNLSEVKKLLTDLSVTDSSEIFKIIEQLEVK 285


RESULT 10
Q8W1R9
ID   Q8W1R9      PRELIMINARY;      PRT;   383 AA.
AC   Q8W1R9;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Geranylgeranyl diphosphate synthase.
OS   Abies grandis (Grand fir).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
OX   NCBI_TaxID=46611;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21671398; PubMed=11733504;
RA   Burke C., Croteau R.;
RT   "Interaction with the Small Subunit of Geranyl Diphosphate Synthase
RT   Modifies the Chain Length Specificity of Geranylgeranyl Diphosphate
RT   Synthase to Produce Geranyl Diphosphate.";
RL   J. Biol. Chem. 277:3141-3149(2002).
DR   EMBL; AF425235; AAL17614.2; -.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; UNKNOWN_1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; UNKNOWN_1.
SQ   SEQUENCE   383 AA;  42069 MW;  EA8D3A015B51CD31 CRC64;

  Query Match             36.9%;  Score 535.5;  DB 10;  Length 383;
  Best Local Similarity   44.9%;  Pred. No. 5.1e-33;
  Matches  122;  Conservative   48;  Mismatches   83;  Indels   19;  Gaps    6;

Qy          2 TNLP---MNKLID----EVNNELSVAINKSV-------MDTQLEESMLYSLNAGGKRIRP 47
              ||||      |:|:    |     ::|:|:::           :: |:| ||| |||||:||
Db         76 TNLPEKVKEKVIEFDFKEYLRSKAMAVNEALDRAVPLRYPERIHEAMRYSLLAGGKRVRP 135

Qy         48 VLLLLTLDSLNTEYELGMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWT 107
              || :   : :     |: | :| |:||||| |||||||| |||||:|||| |||||:|| |
Db        136 VLCISACELVGGTEEVAMPTACAMEMIHTMSLIHDDLPCMDNDDFRRGKPTNHKVFGEGT 195

Qy        108 AILAGDALLTKAFELI---SSDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQ- 163
              |||||||||: ||| |   :|    || : ::|:  |    |  |:|||: |: |||
Db        196 AILAGDALLSFAFEHIAVSTSKSVGTDRI-LRVVSELGRTIGSQGLVGGQVADITSEGDA 254

Qy        164 PIDLETLEMIHKTKTGALLTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDC 223
               :||:||| ||  ||  ||  :||  | |:  |   | :: |:  :|::||: ||:||
Db        255 SVDLDTLEWIHIHKTAVLLECSVMCGAIISGASDNEIERIQRYARSVGLLFQVVDDILDV 314

Qy        224 YGDEAKLGKKVGSDLENNKSTYVSLLGKDGAE 255
                 :|||. | || ::|:||  |:| : |:
Db        315 TKSSKELGKTAGKDLISDKATYPKLMGLEKAK 346


RESULT 11
Q8Z085
ID   Q8Z085      PRELIMINARY;      PRT;   309 AA.
AC   Q8Z085;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Geranylgeranyl diphosphate synthase.
GN   ALR0213.
OS   Anabaena sp. (strain PCC 7120).
OC   Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX   NCBI_TaxID=103690;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21595285; PubMed=11759840;
RA   Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA   Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA   Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA   Yasuda M., Tabata S.;
RT   "Complete genomic sequence of the filamentous nitrogen-fixing
RT   cyanobacterium Anabaena sp. strain PCC 7120.";
RL   DNA Res. 8:205-213(2001).
DR   EMBL; AP003581; BAB77737.1; -.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
KW   Complete proteome.
SQ   SEQUENCE   309 AA;  33706 MW;  E5E2DA7F1E71B657 CRC64;

  Query Match             36.8%;  Score 534.5;  DB 16;  Length 309;
  Best Local Similarity   47.8%;  Pred. No. 4.5e-33;
  Matches  119;  Conservative   35;  Mismatches   92;  Indels    3;  Gaps    2;

Qy         29 QLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKSAIALEMIHTYSLIHDDLPAMD 88
              :: ||| ||| ||||:||:| | | : :    |: | :| |:||||| ||||||||||
Db         44 KIYESMRYSLLAGGKRLRPILCLATSEMMGGTIEIAMPTACAVEMIHTMSLIHDDLPAMD 103
```

```
Qy          89 NDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELIS--SDDRLTDEVKIKVLQRLSIASG 146
               ||||||| |||||||| |||||| ||  ||| ::  : : :  :  ::|: ||  | |
Db         104 NDDYRRGMLTNHKVYGEDIAILAGDGLLAYAFEFVAIRTPESVPRDRVLQVVARLGRALG 163

Qy         147 HVGMVGGQMLDMQSEGQ-PIDLETLEMIHKTKTGALLTFAVMSAADIANVDDTTKEHLES 205
                 |:||||::|: |||:    |||| ||  .|| |||   |:    :|       : |
Db         164 AAGLVGGQVVDLDSEGKSDTSLETLNFIHNHKTAALLEACVVCGGILAGASSENVQRLSR 223

Qy         206 YSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVSLLGKDGAEDKLTYHRDAA 265
               || ::|: ||| ||:||    : :|||  | ||   | || || | : :  |     :||
Db         224 YSQNIGLAFQIIDDILDITSTQEQLGKTAGKDLLAKKVTYPSLWGIEQSRVKAQQLIEAA 283

Qy         266 VDELTQIDE 274
                 ||   |
Db         284 CTELEPFGE 292


RESULT 12
Q9S5F1
ID   Q9S5F1      PRELIMINARY;      PRT;   307 AA.
AC   Q9S5F1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Geranylgeranyl diphosphate synthase (SelGGPS).
GN   CRTE.
OS   Synechococcus elongatus (Thermosynechococcus elongatus).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX   NCBI_TaxID=32046;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99339254; PubMed=10412909;
RA   Ohto C., Ishida C., Nakane H., Muramatsu M., Nishino T., Obata S.;
RT   "A thermophilic cyanobacterium Synechococcus elongatus has three
RT   different Class I prenyltransferase genes.";
RL   Plant Mol. Biol. 40:307-321(1999).
DR   EMBL; AB016093; BAA82613.1; -.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
SQ   SEQUENCE   307 AA;  33101 MW;  A3F918D38B7FB648 CRC64;

  Query Match             36.8%;  Score 534;  DB 2;  Length 307;
  Best Local Similarity   50.9%;  Pred. No. 4.9e-33;
  Matches  114;  Conservative   32;  Mismatches   76;  Indels    2;  Gaps    2;

           29 QLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKSAIALEMIHTYSLIHDDLPAMD 88
              :: ::| ||| ||||||:||:| | | : :    |: | :| ||||||| ||||||||||
           43 KIYDAMRYSLMAGGKRLRPILCLATCELMGGTVEMAMPTACALEMIHTMSLIHDDLPAMD 102

Qy         89 NDDYRRGKLTNHKVYGEWTAILAGDALLTKAFE-LISSDDRLTDEVKIKVLQRLSIASGH 147
              |||||||| |||||||| |||||| ||  ||| ::     :  | :|:: ||  |
Db        103 NDDYRRGKPTNHKVYGEDIAILAGDGLLAYAFEYVVEQTKNVPAEYLLKIVARLGHAVAA 162

Qy        148 VGMVGGQMLDMQSEGQP-IDLETLEMIHKTKTGALLTFAVMSAADIANVDDTTKEHLESY 206
               |:||||::|:: |||| | ||||  ||  |||||  :|:| | :    ::     |  |
Db        163 TGLVGGQVVDLECEGQPDIGLETLHFIHSHKTGALLEASVVSGALLTGAHESDVARLSRY 222

Qy        207 SYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVSLLG 250
              : ::|: ||| ||:||       ||| || |:   | ||  | |
Db        223 AANIGLAFQIVDDILDITSTRDVLGKTVGKDVAAQKMTYPRLWG 266


RESULT 13
Q97HD2
ID   Q97HD2      PRELIMINARY;      PRT;   289 AA.
AC   Q97HD2;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Predicted geranylgeranyl pyrophosphate synthase.
GN   CAC2080.
OS   Clostridium acetobutylicum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1488;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX   MEDLINE=21359325; PubMed=11466286;
RA   Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA   Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA   Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA   Bennett G.N., Koonin E.V., Smith D.R.;
RT   "Genome sequence and comparative analysis of the solvent-producing
RT   bacterium Clostridium acetobutylicum.";
RL   J. Bacteriol. 183:4823-4838(2001).
DR   EMBL; AE007710; AAK80039.1; -.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
KW   Complete proteome.
SQ   SEQUENCE   289 AA;  32381 MW;  C17BAF2903F04DC2 CRC64;

  Query Match             36.6%;  Score 532;  DB 16;  Length 289;
  Best Local Similarity   44.9%;  Pred. No. 6.4e-33;
  Matches  127;  Conservative   44;  Mismatches  108;  Indels    4;  Gaps    3;

Qy          7 NKLI-DEVNNELSVAI-NKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELG 64
              ||:|  ||   ||     |   : :: ||| |||||||||:||:||| :    | :|:
Db          3 NKVIKKEVEEYLSRYFEGKDNYNKRVYESMNYSLNAGGKRVRPLLLLASYAIYNKDYKEV 62

Qy         65 MKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELIS 124
              :  | |:|||||||||||||| ||||| |||: ||||::|   |:|||| || :|  ::
Db         63 IDIAAAIEMIHTYSLIHDDLPCMDNDDLRRGRPTNHKIFGYSIALLAGDGLLNEAMNIMF 122

Qy        125 SDDRLTDEVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQPIDLETLEMIHKTKTGALLTF 184
                    |  :|     :| |:   ||:|||::|: |||: |: : |  :|| ||| |:
Db        123 KYCIGKGEEALKACLMISKAASSDGMIGGQVVDILSEGKKINEDELRYMHKKTGELIKA 182

Qy        185 AVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKST 244
              ||:| | :        | |  |   ||: |||:||:|:||  ||   :||   |||||:| |
Db        183 AVVSGAILGGAPLHDVELLSQYGDKLGLAFQIEDDILDIIGDTKIMGKTSKSDLENDKCT 242

Qy        245 YVSLLGKDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVDL 287
              ||:| | |  :        |  :| : :|  | ||: | || :|
Db        243 YVTLYGIDKCKSICRELTDECLDIIGKI--QGNTELLKEITEL 283


RESULT 14
P72683
ID   P72683      PRELIMINARY;      PRT;   302 AA.
AC   P72683;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Geranylgeranyl pyrophosphate synthase.
GN   CRTE OR SLR0739.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97061201; PubMed=8905231;
RA   Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA   Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
RA   Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA   Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA   Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT   entire genome and assignment of potential protein-coding regions.";
RL   DNA Res. 3:109-136(1996).
DR   EMBL; D90899; BAA16690.1; -.
```

```
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
KW   Complete proteome.
SQ   SEQUENCE   302 AA;  32373 MW;  A7017EADA8296A91 CRC64;

  Query Match             36.2%;  Score 526.5;  DB 16;  Length 302;
  Best Local Similarity   44.2%;  Pred. No. 1.8e-32;
  Matches  117;  Conservative   49;  Mismatches   94;  Indels    5;  Gaps    3;

Qy          4 LPMNKLIDEVNNELSVAINKSVMDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYEL 63
              | : | : |   : |:|| :    :: |:| ||| |||||:||:| :   :    :  |
Db         16 LQVKKGVVEAALDSSLAIAR---PEKIYEAMRYSLLAGGKRLRPILCITACELCGGDEAL 72

Qy         64 GMKSAIALEMIHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFELI 123
               : :| ||||||| |||||||:|||||:|||| |||||||  |||||| ||  ||| :
Db         73 ALPTACALEMIHTMSLIHDDLPSMDNDDFRRGKPTNHKVYGEDIAILAGDGLLAYAFEYV 132

Qy        124 SSDDRLTD-EVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQ-PIDLETLEMIHKTKTGAL 181
               :    | :  ::|: ||    |  |:||||:||::|||: |  |||  ||  |||||
Db        133 VTHTPQADPQALLQVIARLGRTVGAAGLVGGQVLDLESEGRTDITPETLTFIHTHKTGAL 192

Qy        182 LTFAVMSAADIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENN 241
              |  :|:: | :|      :: |  |: ::|: ||: ||:||    : :|||  | |::
Db        193 LEASVLTGAILAGATGEQQQRLARYAQNIGLAFQVVDDILDITATQEELGKTAGKDVKAQ 252

Qy        242 KSTYVSLLGKDGAEDKLTYHRDAAV 266
              |:|| |||| : :  :     | |:
Db        253 KATYPSLLGLEASRAQAQSLIDQAI 277


RESULT 15
Q9ZPM3
ID   Q9ZPM3      PRELIMINARY;      PRT;   393 AA.
AC   Q9ZPM3;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Geranylgeranyl diphosphate synthase.
OS   Taxus canadensis.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Taxaceae; Taxus.
OX   NCBI_TaxID=88032;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99045387; PubMed=9826430;
RA   Hefner J., Ketchum R.E., Croteau R.;
RT   "Cloning and functional expression of a cDNA encoding geranylgeranyl
RT   diphosphate synthase from Taxus canadensis and assessment of the role
RT   of this prenyltransferase in cells induced for taxol production.";
RL   Arch. Biochem. Biophys. 360:62-74(1998).
DR   EMBL; AF081514; AAD16018.1; -.
DR   InterPro; IPR000092; Polyprenyl_synt.
DR   Pfam; PF00348; polyprenyl_synt; 1.
DR   PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR   PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
SQ   SEQUENCE   393 AA;  42572 MW;  6AE0636F385D0F79 CRC64;

  Query Match             35.8%;  Score 520.5;  DB 10;  Length 393;
  Best Local Similarity   41.3%;  Pred. No. 7.5e-32;
  Matches  114;  Conservative   53;  Mismatches  102;  Indels    7;  Gaps    4;

Qy         17 LSVAINKSV---MDTQLEESMLYSLNAGGKRIRPVLLLLTLDSLNTEYELGMKSAIALEM 73
              :  |::|::      :: ||| ||| |||||:|| | :   : :    :| | :| |:||
Db        112 VDAALDKAIPLEYPEKIHESMRYSLLAGGKRVRPALCIAACELVGGSQDLAMPTACAMEM 171

Qy         74 IHTYSLIHDDLPAMDNDDYRRGKLTNHKVYGEWTAILAGDALLTKAFE--LISSDDRLTD 131
              ||| |||||||| |||||:|||| |||||:|| ||:||||||: |||   :::   :
Db        172 IHTMSLIHDDLPCMDNDDFRRGKPTNHKVFGEDTAVLAGDALLSFAFEHIAVATSKTVPS 231

Qy        132 EVKIKVLQRLSIASGHVGMVGGQMLDMQSEGQP-IDLETLEMIHKTKTGALLTFAVMSAA 190
              :  ::|:  |    |  |:||||::|: |||   :||:||| ||  ||  ||  :|:|
Db        232 DRTLRVISELGKTIGSQGLVGGQVVDITSEGDANVDLKTLEWIHIHKTAVLLECSVVSGG 291

Qy        191 DIANVDDTTKEHLESYSYHLGMMFQIKDDLLDCYGDEAKLGKKVGSDLENNKSTYVSLLG 250
              :    :     :  |:  :|::||: ||:||      :|||  | ||  :|:||  |:|
Db        292 ILGGATEDEIARIRRYARCVGLLFQVVDDILDVTKSSEELGKTAGKDLLTDKATYPKLMG 351

Qy        251 KDGAEDKLTYHRDAAVDELTQIDEQFNTKHLLEIVD 286
              : |::      | :||:  | |    || : |
Db        352 LEKAKEFAAELATRAKEELSSFD-QIKAAPLLGLAD 386


Search completed: May 29, 2003, 11:59:52
Job time : 36 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 30, 2003, 19:36:06 ; Search time 2818 Seconds
(without alignments)
2963.983 Million cell updates/sec

Title: US-09-925-637-64
Perfect score: 1453
Sequence: 1 MTNLPMNKLIDEVNNELSVA.....ELTQIDEQFNTKHLLEIVDL 287

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

arched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09925637/runat_23052003_174817_23423/app_query.fasta_1.455
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09925637_@CGN_1_1_2496_@runat_23052003_174817_23423 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	1448	99.7	333750	1	AP004827	AP004827 Staphyloc
c	2	1444	99.4	301550	1	AP003134	AP003134 Staphyloc
c	3	1444	99.4	346900	1	AP003362	AP003362 Staphyloc
	4	1436	98.8	882	6	AR106466	AR106466 Sequence
	5	1015	69.9	3444	1	AF270104	AF270104 Staphyloc
	6	1015	69.9	3444	6	AX145422	AX145422 Sequence
c	7	1015	69.9	4045	1	AF269889	AF269889 Staphyloc
c	8	1015	69.9	4045	6	AX145207	AX145207 Sequence
	9	943	64.9	783	6	AX141501	AX141501 Sequence
	10	702.5	48.3	250050	1	AL591978	AL591978 Listeria
	11	695.5	47.9	5635	6	AX416832	AX416832 Sequence
c	12	686.5	47.2	195269	6	AX417035	AX417035 Sequence
	13	686.5	47.2	333050	1	AL596168	AL596168 Listeria
	14	686.5	47.2	349980	6	AX417044	AX417044 Sequence
	15	657	45.2	894	6	E11868	E11868 Mutated DNA
	16	656	45.1	894	6	E11869	E11869 Mutated DNA
	17	654	45.0	894	6	E54864	E54864 Process for
	18	653	44.9	894	6	E11866	E11866 Mutated DNA
	19	650	44.7	894	6	AR148386	AR148386 Sequence
	20	650	44.7	894	6	E05337	E05337 DNA encodin
	21	650	44.7	894	6	E11867	E11867 Mutated DNA
	22	650	44.7	894	6	E14762	E14762 gDNA encodi
	23	650	44.7	894	6	E16095	E16095 DNA encodin
	24	650	44.7	894	6	E27521	E27521 Geranyl dip
	25	650	44.7	1260	1	BACFDPS	D13293 B. stearoth
	26	649	44.7	1080	1	AB003187	AB003187 Micrococc
	27	645	44.4	894	6	E11865	E11865 Mutated DNA
	28	644	44.3	894	6	E27520	E27520 Geranyl dip
c	29	638	43.9	300950	1	AP001516	AP001516 Bacillus
	30	635	43.7	811	6	AX432600	AX432600 Sequence
c	31	632.5	43.5	218470	1	BSUB0013	Z99116 Bacillus su
	32	632.5	43.5	282700	1	BACJH642	D84432 Bacillus su
	33	593	40.8	2268	1	LMO012349	AJ012349 Listeria
c	34	576	39.6	11055	1	AE010637	AE010637 Fusobacte
c	35	556.5	38.3	13536	1	AE010066	AE010066 Streptoco
c	36	555.5	38.2	50463	1	AE014159	AE014159 Streptoco
c	37	551.5	38.0	10165	1	AE006583	AE006583 Streptoco
	38	551	37.9	876	6	AR112484	AR112484 Sequence
c	39	551	37.9	296750	1	AP003191	AP003191 Clostridi
c	40	549	37.8	10011	6	BD003756	BD003756 Polynucle
c	41	549	37.8	11999	1	AE007421	AE007421 Streptoco
c	42	549	37.8	19059	2	SPNEU1917	AL449939 Streptoco
c	43	547	37.6	10698	1	AE008481	AE008481 Streptoco
	44	540.5	37.2	10029	1	AE006320	AE006320 Lactococc
	45	535.5	36.9	1548	8	AF425235	AF425235 Abies gra

ALIGNMENTS

RESULT 1

```
AP004827/c
LOCUS       AP004827              333750 bp    DNA     linear   BCT 02-JUL-2002
DEFINITION  Staphylococcus aureus subsp. aureus MW2 DNA, complete genome,
            strain:MW2, section 6/10.
ACCESSION   AP004827 BA000033
VERSION     AP004827.1  GI:21204509
KEYWORDS    .
SOURCE      Staphylococcus aureus subsp. aureus MW2 (strain:MW2) DNA.
  ORGANISM  Staphylococcus aureus subsp. aureus MW2
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Baba,T., Takeuchi,F., Kuroda,M., Yuzawa,H., Aoki,K., Oguchi,A.,
            Nagai,Y., Iwama,N., Asano,K., Naimi,T., Kuroda,H., Cui,L.,
            Yamamoto,K. and Hiramatsu,K.
  TITLE     Genome and virulence determinants of high virulence
            community-acquired MRSA
  JOURNAL   Lancet 359 (9320), 1819-1827 (2002)
  MEDLINE   22040717
   PUBMED   12044378
REFERENCE   2  (bases 1 to 333750)
  AUTHORS   Aoki,K., Oguchi,A., Nagai,Y., Asano,K., Iwama,N., Baba,T.,
            Kuroda,M., Hiramatsu,K. and Kikuchi,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAR-2002) Akio Oguchi, National Institute of
            Technology and Evaluation, Biotechnology Center; 2Chome 49-10
            Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
            (E-mail:oguchi@nite.go.jp, URL:http://www.bio.nite.go.jp/,
            Tel:81-3-3481-8423, Fax:81-3-3481-8424)
FEATURES             Location/Qualifiers
     source          1. .333750
                     /organism="Staphylococcus aureus subsp. aureus MW2"
                     /strain="MW2"
                     /db_xref="taxon:196620"
     gene            complement(278. .619)
                     /gene="MW1341"
     CDS             complement(278. .619)
                     /gene="MW1341"
                     /note="ORFID:MW1341"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAB95206.1"
                     /db_xref="GI:21204510"
                     /translation="MTNITKEVFDNLEQEIDVFAKNKTLGSSEAKPYLDEYHSKIIDY
                     FKQVNDITGNIDFDKLNQYPVVPMNFKERYDYMIERKYHFMGYRQMKTFKTELIKMNA
                     SYQTRLKNKQV"
     gene            complement(624. .1283)
                     /gene="nth"
     CDS             complement(624. .1283)
                     /gene="nth"
                     /note="ORFID:MW1342"
                     /codon_start=1
                     /transl_table=11
                     /product="endonuclease-like protein"
                     /protein_id="BAB95207.1"
                     /db_xref="GI:21204511"
                     /translation="MVSKKKALEMIDVIANMFPDAECELKHDNPFELTIAVLLSAQCT
                     DVLVNRVTTELFKKYKTPEDYLAVSDEELMNDIRSIGLYRNKAKNIKKLCQSLIDQFN
                     GEIPQTHKELESLAGVGRKTANVVMSVAFDEPSLAVDTHVERVSKRLGINRWKDNVRQ
                     VEDRLCSVIPRDRWNRSHHQLIFFGRYHCLARKPKCDICPLLEDCREGQKRYKASLKE
                     A"
     gene            complement(1273. .1959)
                     /gene="MW1343"
     CDS             complement(1273. .1959)
                     /gene="MW1343"
                     /note="ORFID:MW1343
                     hypothetical protein, similar to chromosome replication
                     initiation protein dnaD"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB95208.1"
                     /db_xref="GI:21204512"
                     /translation="MDKYQLKARPVVIRRELLDHYSDLGLDEQDLVILLKLIYASETS
                     NKQPSIELLQKGSTMQPRDITMVIQNLIQRELLELQVQKDEEGRFTEYMNLDPFFEKL
                     SHILKQQSMETKEQNSKEKFKQLFRVLEDTFARPLSPYEIETLNQWIDVDKHDSAIIQ
                     AALDEANSLNKLSFKYMDRILLNWKKNNVKTIDDSRKIREKFNKPKMTHTVKTVPKFD
                     WLNGENLDGK"
     gene            complement(2287. .3579)
                     /gene="asnS"
     CDS             complement(2287. .3579)
                     /gene="asnS"
                     /note="ORFID:MW1344"
                     /codon_start=1
                     /transl_table=11
                     /product="asparaginyl-tRNA synthetase"
                     /protein_id="BAB95209.1"
                     /db_xref="GI:21204513"
                     /translation="MKTTIKQAKDHLNQDVTIGAWLTNKRSSGKIAFLQLRDGTGFMQ
                     GVVVKSEVDEEVFKLAKEITQESSLYVTGTITEDNRSDLGYEMQVKSIEVISEAHDYP
                     ITPKNHGTEFLMDHRHLWLRSKKQHAVMKIRNEVIRATYEFFNKDGFTKVDPPILTAS
                     APEGTSELFHTKYFDQDAFLSQSGQLYLEAAAMAHGKVFSFGPTFRAEKSKTRRHLIE
                     FWMIEGEMAFTNHAESLEIQEQYVTHVVKSVLENCKLELKILERDTSKLEKVATPFPR
                     ISYDDAIEFLKSEGFDDIEWGEDFGAPHETAIANHYDLPVFITNYPTKIKPFYMQPNP
                     ENEETVLCADLIAPEGYGEIIGGSERVDDLELLEQRVKEHGLDEEAYSYYLDLRRYGS
                     VPHCGFGLGLERTVAWISGVEHVRETAPFPRLLNRLYP"
     gene            complement(3901. .6594)
                     /gene="dinG"
     CDS             complement(3901. .6594)
                     /gene="dinG"
                     /note="ORFID:MW1345"
                     /codon_start=1
                     /transl_table=11
                     /product="probable ATP-dependent DNA helicase dinG"
                     /protein_id="BAB95210.1"
                     /db_xref="GI:21204514"
                     /translation="MGMATYAVVDLETTGNQLDFDDIIQIGITFVRNNQIIDTYHSMI
                     RTNLEIPPFIQALTSIEENMLQQAPYFNQVAQEIYDKIKDCIFVAHNVDFDLNFIKKA
                     FKDCNIQYRPKKVIDTLEIFKIAFPTDKSYQLSELAEAHGITLANAHRADEDAATTAK
                     LMILAFEKFEKLPLDTLKQLYYLSKQLKYDLYDIFFEMVRQYDAKPLDKSYEKFEQII
                     YRKQVDFKKPTTNYNGSLKSLYSKAVDQLGLTYRPQQLYLAETILDQLMHSEKAMIEA
                     SLGSGKSLAYLLAALMYNIETGKHVMISTNTKLLQSQLLEKDIPAMNEALNFKINALL
                     IKSKSDYISLGLISQILKDDTSNYEVNILKMQLLIWITETPSGDIQELNLKGGQKMYF
                     DQKIETYVPARHDVHYYNFIKRNAQNIQIGITNHAHLIHSDVENSIYQLFDDCIVDEA
                     HRLPDYALNQVTNELSYADIKYQLGLIGKNENEKLLKAIDQLEKQRILEKLDIAPIDI
                     FGLKASMNEIHELNEQLFSTIFTIINDSDVYDDDIHRFHNVFTFETKDILKDLHAIID
                     KLNKTLEIFNGISHKTVKSLRKQLLYLKDKFKNIEQSLKAGHTSFISIKNLSQKSTIR
                     LYVKDYAVKDVLTKQVLEKFKSLIFISGTLKFNHSFEAFKQLFNKDVHFNTFEVNTSL
                     QSAKNTSVFIPSDVASYQYKNIDEYVASIVSYIIEYTTITSSKCLVLFTSYKMMHMVQ
                     DMLNELPEFEDYVVLTQQQNQNYKIVQQFNNFDKAILLGTSTFFEGFDFQANGIKCVM
                     IAKLPFMNKHNAKYWLMDSEFTSTFKEYVLPDAVTRFRQGLGRLIRSENDRGIIVSFD
                     DRLINSNYKNFFEQTLENYRQKKGDIQQFGKLLRQIQKKKK"
     gene            complement(6618. .7589)
                     /gene="MW1346"
     CDS             complement(6618. .7589)
                     /gene="MW1346"
                     /note="ORFID:MW1346"
                     /codon_start=1
                     /transl_table=11
                     /product="truncated hypothetical protein"
                     /protein_id="BAB95211.1"
                     /db_xref="GI:21204515"
                     /translation="MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLK
                     LEGCKIDSVNHKGHLLQQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSL
                     VGNQSSFFILSDEQTKGRGRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALG
                     IRDAIQHFSQDEVKVKWPNDIYIDNGKVCGFLTEMVANNDGIEAIICGIGINLTQQLE
                     NFDESIRHRATSVQLHDKNKLDRYQFLERLLQEIEKRYNQFLTLPFSEIREEYIAASN
                     IWNRTLLFTENDKQFKGQAIDLDYDGYLIVRDEAGESHRLISADIDF"
     gene            complement(7576. .8778)
                     /gene="MW1347"
     CDS             complement(7576. .8778)
                     /gene="MW1347"
                     /note="ORFID:MW1347
                     hypothetical protein, similar to poly(A) polymerase"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB95212.1"
```

```
                     /db_xref="GI:21204516"
                     /translation="MDKSLFEQARPILEQIQDNGFEAYYVGGSVRDYVMGRNIHDIDI
                     TTSATPDEIESIFSHTIPVGKEHGTINVVFNDENYEVTTFRAEEDYVDHRRPSGVTFV
                     RDLYEDLQRRDFTMNAIAMDTAYKLYDYFDGQQDINNRIIRTVGIAEERFQEDALRMI
                     RCLRFQSQLSFDIATETFEAMRIQMADIKFLSIERIVIELTKLMRGINVEKSFNHLKS
                     LKAFNYMPYFEHLDMNQINVTEAIDLELLIAIVSVKFDINYSLKPLKLSNRQVKDINQ
                     YIQIMNALPSIITKEQLKMFVYDYDTHLIKNVMVAADVIKANDIQGHEPLIVNLQTID
                     ETLHRLPMHNRKDMMVNGGVLMAHLNAKSGPWLKDVLRQIEIAIVTGKVSNEETEILK
                     WVDNHVKI"
     gene            complement(8783. .9925)
                     /gene="MW1348"
     CDS             complement(8783. .9925)
                     /gene="MW1348"
                     /note="ORFID:MW1348
                     hypothetical protein, similar to lipopolysaccharide
                     biosynthesis-related pr homolog"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB95213.1"
                     /db_xref="GI:21204517"
                     /translation="MKIGITCYPSMGGSGIIATELGIKLAERGHEVHFITSNIPFRIR
                     KPLPNMIFHQVEVNQYAVFQYPPYDITLSTKIAEVIKEYGLDLLHMHYAVPHAICGIL
                     AREMSGKDIKIMTTLHGTDITVLGYDHSLQGAIKFGIEKSDIVTSVSKSLAQETHEII
                     ETNKEIIPIYNFVRENEFPTKHNTALKSQFGIAPDEKVLIHVSNFRQVKRIDTIIETF
                     AKVREKIPSKLILLGDGPELVPMRQLTKELNVEEDVLFLGKQDCVSEFYQLSDLVLLL
                     SEKESFGLTLLEAMKTGVVPIGSNAGGIKEVIKHGETGFVVDVGDCDSASDYAIRLLE
                     DKVLYNKLQKNMLADIAERFGSELITDQYEYYYQKMLNEHNKSKGE"
     gene            complement(10167. .10484)
                     /gene="MW1349"
     CDS             complement(10167. .10484)
                     /gene="MW1349"
                     /note="ORFID:MW1349"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB95214.1"
                     /db_xref="GI:21204518"
                     /translation="MKSMVEMQREVDEYIGQFKTGYFSPLANLARLTEEVGELAREIN
                     HTYGEKKKKDSEEANTIKAELGDNLFVLLCLANSMGIDMTESFNETMEKFNTRDKNRF
                     ERK"
     gene            complement(10820. .11518)
                     /gene="MW1350"
     CDS             complement(10820. .11518)
                     /gene="MW1350"
                     /note="ORFID:MW1350"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB95215.1"
                     /db_xref="GI:21204519"
                     /translation="MSSIFTMIIYFVILMVLPLWAQHKVKSNYEKYSQVRSTSGKTGR
                     EVALEILHANGIYDVEVVKGEGFLTDHYDPKKKVVSLSPANYDRPSVAGTAIAAHEVG
                     HAIQDHQGYWFLRFRAALVPVANLGSSLSYMIIMLGIILTAIGSAFGSTALWIGAGLM
                     SLAVLFSIVTLPVEFDASSRAMKQITALNIVNEKEYKHARKVLSAAAMTYVAATAVAV
                     AELVRIILLARSSD"

Alignment Scores:
Pred. No.:              6.72e-113     Length:          333750
Score:                  1448.00       Matches:         286
Percent Similarity:     99.65%        Conservative:    0
Best Local Similarity:  99.65%        Mismatches:      1
Query Match:            99.66%        Indels:          0
DB:                     1             Gaps:            0

US-09-925-637-64 (1-287) x AP004827 (1-333750)

Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118374 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 118315

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118314 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 118255

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118254 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 118195

Qy         61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118194 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 118135

Qy         81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118134 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 118075

Qy        101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118074 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 118015

Qy        121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118014 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 117955

Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117954 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 117895

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117894 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 117835

Qy        181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117834 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAA 117775

Qy        201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117774 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 117715

Qy        221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117714 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 117655

Qy        241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117654 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 117595

Qy        261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117594 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 117535

Qy        281 LeuLeuGluIleValAspLeu 287
              |||||||||||||||||||||
Db     117534 TTATTAGAAATCGTTGATTTA 117514

RESULT 2
AP003134/c
LOCUS       AP003134               301550 bp    DNA     linear   BCT 02-JUL-2002
DEFINITION  Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
            genome, section 6/10.
ACCESSION   AP003134 BA000018
VERSION     AP003134.2  GI:14349226
KEYWORDS    .
SOURCE      Staphylococcus aureus subsp. aureus N315 (sub_species:aureus N315,
            strain:N315) DNA.
  ORGANISM  Staphylococcus aureus subsp. aureus N315
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
            Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M.,
            Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
            Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
            Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
            Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
            Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
```

```
  TITLE     Whole genome sequencing of meticillin-resistant Staphylococcus
            aureus
  JOURNAL   Lancet 357 (9264), 1225-1240 (2001)
  MEDLINE   21311952
   PUBMED   11418146
REFERENCE   2  (bases 1 to 301550)
  AUTHORS   Aoki,K., Oguchi,A., Hosoyama,A., Nagai,Y., Kuroda,M., Hiramatsu,K.
            and Kikuchi,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JAN-2001) Akio Oguchi, National Institute of
            Technology and Evaluation, Biotechnology Center; 2Chome 49-10
            Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
            (E-mail:oguchi@nite.go.jp, URL:http://www.bio.nite.go.jp/,
            Tel:81-3-3481-8423, Fax:81-3-3481-8424)
COMMENT     On Jun 12, 2001 this sequence version replaced gi:13701258.
FEATURES             Location/Qualifiers
     source          1. .301550
                     /organism="Staphylococcus aureus subsp. aureus N315"
                     /strain="N315"
                     /sub_species="aureus N315"
                     /db_xref="taxon:158879"
     gene            complement(249. .947)
                     /gene="SA1293"
     CDS             complement(249. .947)
                     /gene="SA1293"
                     /note="ORFID:SA1293"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB42553.1"
                     /db_xref="GI:13701259"
                     /translation="MSSIFTMIIYFVILMVLPLWAQHKVKSNYEKYSQVRSTSGKTGR
                     EVALEILHANGIYDVEVVKGEGFLTDHYDPKKKVVSLSPANYDRPSVAGTAIAAHEVG
                     HAIQDHQGYWFLRFRAALVPVANLGSSLSYMIIMLGIILTAIGSAFGSTALWIGAGLM
                     SLAVLFSIVTLPVEFDASSRAMKQITALNIVNEKEYKHARKVLSAAAMTYVAATAVAV
                     AELVRIILIARSSD"
     gene            complement(1001. .1588)
                     /gene="SA1294"
     CDS             complement(1001. .1588)
                     /gene="SA1294"
                     /note="ORFID:SA1294"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB42554.1"
                     /db_xref="GI:13701260"
                     /translation="MTINTFWQYTLYQRSWLMMLLICNILGMIYGYIWYGEQLSHTPW
                     HFKIFVPDSPTAILYLVISISLILIQKQNSIIEALAFVTLFKYGIWAVIMNILFIIEQ
                     GDITVNGLVLMFSHSIMAVQAIYFYPRFKRSMIGISVAMIWVFLNDYIDYFHLQFPYY
                     DFITTHVWQIGVLSCCLSVFGLLLYIELNKLLKCK"
     gene            complement(1578. .2153)
                     /gene="SA1295"
     CDS             complement(1578. .2153)
                     /gene="SA1295"
                     /note="ORFID:SA1295"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB42555.1"
                     /db_xref="GI:13701261"
                     /translation="MSETLNQIKESFIEYLLFQYRFKSRIAVWVLNYIKVNEAKLANI
                     HFVDTKINHHETLEIAEVGSHASAIQFTKRNIKLMNTNEIFDYIANHNCAFDIQIHFA
                     NVSKREQRLDDLIVAQLTESPSYQTYLHDLNSMAIDRHKHALLIDYLLHNIDLSLQMN
                     EKQRFYQLTQILNTLKLVNKHNQFEDLADDD"
     gene            complement(2167. .3411)
                     /gene="SA1296"
     CDS             complement(2167. .3411)
                     /gene="SA1296"
                     /note="ORFID:SA1296"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB42556.1"
                     /db_xref="GI:13701262"
                     /translation="MEDIYKLIDDINLQKLENLDSRVNEAITTDNDDALFILGETLYN
                     FGLMPQGLEVFRVLYHKYPDESELLIYFIEGLMSENQTDEALEYLSYVEPSPEKLMLE
                     ADLYQQINMMEVAIDKLQEALELEPNDPIIHFALAEMLYYDGQYLRATSEYETVLETG
                     EYQVNGVNLFSRMADCSLQSGNYSDAIRLYDEINEDEMTSEDYLKKAISYDKNDITQE
                     AIKIMTTLLSKDPDYIQGYLYLQSLYENEKNYPDAIETGKEGLRLSQFYKELMYTTGC
                     LEIEHGDANEGVLLLKQALEVDNAYQEPLLILSDLYRNEEDYEAIIELLTYVDEEDLD
                     PTFMWHLAFAYGQEERDKEAQHFFELAYPTMKTNIDFMSDYYFYLIEIGQKEQAITIL
                     NQLLELEPSNENWHDELSRLQS"
     gene            complement(3418. .4716)
                     /gene="aroA"
     CDS             complement(3418. .4716)
                     /gene="aroA"
                     /note="ORFID:SA1297"
                     /codon_start=1
                     /transl_table=11
                     /product="3-phosphoshikimate 1-carboxyvinyltransferase"
                     /protein_id="BAB42557.1"
                     /db_xref="GI:13701263"
                     /translation="MVSEQIIDISGPLKGEIEVPGDKSMTHRAIMLASLAEGTSNIYK
                     PLLGEDCRRTMDIFRLLGVDIKEDEDKLVVNSPGYKAFKTPHQVLYTGNSGTTTRLLA
                     GLLSGLGIESVLSGDVSIGKRPMDRVLRPLKLMDANIEGIEDNYTPLIIKPSVIKGIN
                     YQMEVASAQVKSAILFASLFSNDTTVIKELDVSRNHTETMFRHFNIPIEAERLSITTT
                     PDAIQHIKPADFHVPGDISSAAFFIVAALITPESDVTIHNVGINPTRSGIIDIVEKMG
                     GNIQLFNQTTGAEPTASIRIQYTPMLQPITIEGELVPKAIDELPVIALLCTQAVGTST
                     IKDAEELKVKETNRIDTTADMLNLLGFELQPTNDGLIIHPSEFKTNATVDSLTDHRIG
                     MMLAVASLLSSEPVKIKQFDAVNVSFPGFLPKLKLLENEG"
     gene            complement(4726. .5790)
                     /gene="aroB"
     CDS             complement(4726. .5790)
                     /gene="aroB"
                     /note="ORFID:SA1298"
                     /codon_start=1
                     /transl_table=11
                     /product="3-dehydroquinate synthase"
                     /protein_id="BAB42558.1"
                     /db_xref="GI:13701264"
                     /translation="MKLQTTYPSNNYPIFVEHGAIDHISTYIDQFDQSFILIDEHVNQ
                     YFADKFDDILSYENVHKVIIPAGEKTKTFEQYQETLEYILSHHVTRNTAIIAVGGGAT
                     GDFAGFVAATLLRGVHFIQVPTTILAHDSSVGGKVGINSKQGKNLIGAFYRPTAVIYD
                     LDFLKTLPFEQILSGYAEVYKHALLNGESTTQEIEQHFKDREILQSLNGMDKYIAKGI
                     ETKLDIVVADEKEQGVRKFLNLGHTFGHAVEYNHKIAHGHAVMIGIIYQFIVANILFN
                     SNHDIQHYINYLTKLGYPLETITDIDFETIYQYMLSDKKNDKQGVQMVLIKHFGDIVV
                     QHIDQTTLQHACEQLKTYFK"
     gene            complement(5816. .6982)
                     /gene="aroC"
     CDS             complement(5816. .6982)
                     /gene="aroC"
                     /note="ORFID:SA1299"
                     /codon_start=1
                     /transl_table=11
                     /product="chorismate synthase"
                     /protein_id="BAB42559.1"
                     /db_xref="GI:13701265"
                     /translation="MRYLTSGESHGPQLTVIVEGIPANLEIKVEDINKEMFKRQGGYG
                     RGRRMQIEKDTVEIVSGVRNGYTLGSPITMVVTNDDFTHWRKIMGAAPISEEERENMK
                     RTITKPRPGHADLVGGMKYNHRDLRNVLERSSARETAARVAVGALCKVLLQQLDIDIY
                     SRVVEIGGIKDKDFYDSETFKANLDRNDVRVIDDSIAQAMRDKIDEAKNEGDSIGGVV
                     QVVVENMPVGVGSYVHYDRKLDGKIAQGVVSINAFKGVSFGEGFKAAEKPGSEIQDEI
                     LYNSEIGYYRGSNHLGGLEGGMSNGMPIIVNGVMKPIPTLYKPLNSVDINTKEDFKAT
                     IERSDSCAVPAASIVCEHVVAFEIAKALLEEFQSNHIEQLQQQIADRRQLNVEF"
     gene            complement(7368. .7568)
                     /gene="SA1300"
     CDS             complement(7368. .7568)
                     /gene="SA1300"
                     /note="ORFID:SA1300"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAB42560.1"
                     /db_xref="GI:13701266"
                     /translation="MKNIISIILGILMFLKLMELLYGAIFLDKPLNPITKIIFILTLI
                     YIFYVLVKELIIFLKSKYNKSA"
```

```
     gene            complement(7597. .7722)
                     /gene="SAS045"
     CDS             complement(7597. .7722)
                     /gene="SAS045"
                     /note="ORFID:SAS045"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAB42561.1"
                     /db_xref="GI:13701267"
                     /translation="MWLVIMIIHNKARFLCCSKLLEKFYSFLVTRIKKRGIFFIK"
     gene            complement(7777. .8226)
                     /gene="ndk"
     CDS             complement(7777. .8226)
                     /gene="ndk"
                     /note="ORFID:SA1301"
                     /codon_start=1
                     /transl_table=11
                     /product="nucleoside diphosphate kinase"
                     /protein_id="BAB42562.1"
                     /db_xref="GI:13701268"
                     /translation="MERTFLMIKPDAVQRNLIGEVISRIERKGLKLVGGKLMQVPMEL
                     AETHYGEHQGKPFYNDLISFITSAPVFAMVVEGEDAVNVSRHIIGSTNPSEASPGSIR
                     GDLGLTVGRNIIHGSDSLESAEREINLWFNENEITSYASPRDAWLYE"
     gene            complement(8318. .9277)
                     /gene="gerCC"
     CDS             complement(8318. .9277)
                     /gene="gerCC"
                     /note="ORFID:SA1302"
                     /codon_start=1
                     /transl_table=11
                     /product="heptaprenyl diphosphate syntase component II"
                     /protein_id="BAB42563.1"
                     /db_xref="GI:13701269"
                     /translation="MAKLNMNNEIKKVEQRLEKAIKSKDSVLEQASLHLLSSGGKRVR
                     PAFVILSSQFGKDEQTSEQTYQVAVALELIHMATLVHDDVIDKSDKRRGKLTISKKWD

Alignment Scores:
Pred. No.:              1.31e-112       Length:         301550
Score:                  1444.00         Matches:        285
Percent Similarity:     99.65%          Conservative:   1
Best Local Similarity:  99.30%          Mismatches:     1
Query Match:            99.38%          Indels:         0
DB:                     1               Gaps:           0

US-09-925-637-64 (1-287) x AP003134 (1-301550)

Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        64130 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 64071

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64070 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGCATGTTGTATTCATTAAATGCT 64011

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64010 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 63951

Qy         61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
              ||||||||||:::||||||||||||||||||||||||||||||||||||||||||||||
Db      63950 TATGAGTTAGGTGTGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 63891

Qy         81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63890 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCAT 63831

Qy        101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63830 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 63771

Qy        121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db      63770 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 63711

Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63710 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 63651

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63650 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 63591

Qy        181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63590 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAA 63531

Qy        201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63530 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 63471

Qy        221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63470 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 63411

Qy        241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63410 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 63351

Qy        261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63350 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 63291

Qy        281 LeuLeuGluIleValAspLeu 287
              |||||||||||||||||||||
Db      63290 TTATTAGAAATCGTTGATTTA 63270

RESULT 3
AP003362/c
LOCUS       AP003362               346900 bp    DNA     linear   BCT 07-FEB-2002
DEFINITION  Staphylococcus aureus subsp. aureus Mu50 genomic DNA, complete
            sequence, section 5/9.
ACCESSION   AP003362 BA000017
VERSION     AP003362.2  GI:14247083
KEYWORDS    .
SOURCE      Staphylococcus aureus subsp. aureus Mu50 (sub_species:aureus Mu50,
            strain:Mu50) DNA.
  ORGANISM  Staphylococcus aureus subsp. aureus Mu50
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
            Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M.,
            Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
            Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
            Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
            Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
            Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
  TITLE     Whole genome sequencing of meticillin-resistant Staphylococcus
            aureus
  JOURNAL   Lancet 357 (9264), 1225-1240 (2001)
  MEDLINE   21311952
REFERENCE   2  (bases 1 to 346900)
  AUTHORS   Ohta,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Toshiko Ohta, University of Tsukuba College
            of Medical Technology and Nursing, Department of Medical
            Technology; 1-1-1 Ten-noodai, Tsukuba, Ibaraki 305-8577, Japan
            (E-mail:tohta@sakura.cc.tsukuba.ac.jp, Tel:81-298-53-3454,
            Fax:81-298-53-3454)
COMMENT     On May 29, 2001 this sequence version replaced gi:13875626.
FEATURES             Location/Qualifiers
     source          1. .346900
                     /organism="Staphylococcus aureus subsp. aureus Mu50"
                     /strain="Mu50"
                     /sub_species="aureus Mu50"
                     /db_xref="taxon:158878"
```

```
gene            12. .263
                /gene="SAV1313"
CDS             12. .263
                /gene="SAV1313"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB57475.1"
                /db_xref="GI:14247084"
                /translation="MLTLIKLKEDEQVIIYEYIPEDDISNGKGSVTFNKKDAEVIDFS
                LSEIENEEYFMLYRNKSFSVVRDFIEKQEFPENYKIAWY"
gene            987. .1181
                /gene="SAV1314"
CDS             987. .1181
                /gene="SAV1314"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB57476.1"
                /db_xref="GI:14247085"
                /translation="MNLQSYIGKLVKLTLTNNKILIGKVIDFDDKVDNFDGYNSIEID
                TGRITYDISENKIKTIVLQK"
gene            complement(1725. .2750)
                /gene="SAV1315"
CDS             complement(1725. .2750)
                /gene="SAV1315"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="BAB57477.1"
                /db_xref="GI:14247086"
                /translation="MISFENDYLEGAHEKVLKRLVDTNLVQASGYGFDQFTAQAIEKI
                KDTIDCPNATIRFLVGGTQTNQVVINSMLESYEGVISADTGHVAVHEGGAIEYSGHKV
                ITIPSKEGKVSASDVETYMETFKSDFKKDHMVFPGMVYISHPTEYGTLYSKSELEELC
                KVCKQYQLPLFMDGARLGYGLMSDQSDMTIKDIAKYCDVFYIGGTKIGALCGEAIVFT
                KNNEPKQFTTRIKHHGALLAKGRLTGIQFLELFTDNLYFNISRHAIEMANKMKDGFKN
                KGYRLYFDSPTNQQFFILSNEKIAELEQKVKFAVWEKYDDQHRVVRFATSWATTEENL
                NKLLELI"
gene            3020. .3217
                /gene="SAV1316"
CDS             3020. .3217
                /gene="SAV1316"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB57478.1"
                /db_xref="GI:14247087"
                /translation="MEQIKLKTFTAETLELLEKNINAFLSSEEATNLKLVNITIKEIE
                ERTFPNNEEEFNAILTLSVNK"
gene            3238. .4755
                /gene="SAV1317"
CDS             3238. .4755
                /gene="SAV1317"
                /codon_start=1
                /transl_table=11
                /product="cardiolipin synthetase homolog"
                /protein_id="BAB57479.1"
                /db_xref="GI:14247088"
                /translation="MFFVPFFIGEKTMRYTFSNDLGTLFTIILAIGFIINLVLAFIII
                FLERNRRTASSTWAWLFVLFVLPLIGFILYLFFGRTVSARKLNKNNGNVLTDFDGLLK
                QQIESFDKGNYGTDNKQVQKHHDLVSMLLMDQDGFLTENNKVDHFIDGNDLYDQVLKD
                IKNAKEYIHLEYYTFALDGLGKRILHALEEKLKQGLEVKILYDDVGSKNVKMANFDHF
                KSLGGEVEAFFASKLPLLNFRMNNRNHRKIIVIDGQLGYVGGFNIGDEYLGLGKLGYW
                RDTHLRIQGDAVDALQLRFILDWNSQAHRPQFEYDVKYFPKKNGPLGNSPIQIAASGP
                ASDWHQIEYGYTKMIMSAKKSVYLQSPYFIPDNSYINAIKIAAKSGVDVHLMIPCKPD
                HPLVYWATFSNASDLLSSGVKIYTYENGFIHSKMCLIDDEIVSVGTANMDFRSFELNF
                EVNAFVYDENLAKDLRVAYEHDITKSKQLTKESYANRPLSVKFKESLAKLVSPIL"
gene            4939. .5436
                /gene="SAV1318"
CDS             4939. .5436
                /gene="SAV1318"
                /note="ATP-binding protein"
                /codon_start=1
                /transl_table=11
                /product="ABC transporter homolog"
                /protein_id="BAB57480.1"
                /db_xref="GI:14247089"
                /translation="MIQISNINKSFNKRCVLKNISFDIEQGKCIALIGKNGAGKSTLI
                DILIGNVNANSGEIFDKDKLLQSENRSIMFQKTMFPDQLKVIEIINLYQSFYENPLPL
                EEIIELTKFDSSQLNQFVNKLSGGQQRLLDFVLSLIGQPQLILLDEPTSTMDIEIREY
                FWSII"
gene            5461. .5838
                /gene="SAV1319"
CDS             5461. .5838
                /gene="SAV1319"
                /note="ATP-binding protein"
                /codon_start=1
                /transl_table=11
                /product="ABC transporter homolog"
                /protein_id="BAB57481.1"
                /db_xref="GI:14247090"
                /translation="MLYTSHYIEEVERMSDKIILIENGEIILNDSTSHIRTNQQSQIT
                LSDEYIRKLKLDKDDLVIQKNHNGTIKIITSNVNDTILYLQQLHINLDDIEIQKVSIV
                DSYFNNKKQRGSNYDTKLLENRI"
gene            5807. .6538
                /gene="SAV1320"
CDS             5807. .6538
                /gene="SAV1320"
                /note="similar to ABC transporter integral membrane
                protein"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB57482.1"
                /db_xref="GI:14247091"
                /translation="MILSYLKIEFKVIMRKKTTLILSILFPVIFYILFTSILELPEDV
                KPKFYKEYMYSMTVYSLLSFSLLTFPLDIINEKQNEWRQRLMVTPFTFTSYYISKVVK
                TMLQFAIAILVIFMVGHFYKGVAMSAVQWLESGIFLWLGASLLITFGILFSLLNDIQK
                TSALANIVTIGLAVLGGLWFPINTFPNWLQHVAHVLPSYHLRKLGVDIASNHHINLIS
                FAIILLYVLGSIIAVYCISHFKRAE"
gene            6542. .7633
                /gene="SAV1321"
CDS             6542. .7633
                /gene="SAV1321"
                /note="similar to two-component sensor histidine kinase"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB57483.1"
                /db_xref="GI:14247092"
                /translation="MKFLKDTSIAEISSILYLIFPIAGIFFNEVYGPKWLYIISVIVF
                SLSYLILVIVNNRLNTLMFYILLIIHYFIICYFVFSVHPMLSLFFFYSAFAVPFTFKN
                NVKKTATNLFILTMIICTIITYLLYNNYFVAMMVYYVVISLIMLDNFKKMKNREYQKE
                IAEKNRHINTLIAEQERHRIGQDLHDTLGHVFASLSLKSELAYKLIDTDVEKVKAELL
                AINKLSRESLNKVREIIDDVKLPSFIEEIDSIRKVLKDADIDFTFENKELAQVLSPTK
                QSMLVMITREAINNVIKHANASKVHGKLKTVNNHKLLLMIEDDGKGIDSDCEVKSISQ
                RVQHLNGTLAVDSTNGTKIIIEISTGGIA"
gene            7630. .8232
                /gene="SAV1322"
CDS             7630. .8232
                /gene="SAV1322"
                /note="similar to two-component response regulator"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB57484.1"
                /db_xref="GI:14247093"
                /translation="MTSLIIAEDQNMLRQAMVQLIKLHGDFEILADTDNGLDAMKLIE
                EYNPNVVILDIEMPGMTGLEVLAEIRKKHLNIKVIIVTTFKRPGYFEKAVVNDVDAYV
                LKERSIEELVETINKVNNGEKEYSATLMTSFFVDKNPLTPKEQIVLREIGNGLSSKEI
                SEKLFLTDGTVRNYTSVIIDKLFADNRFDAWKKANEKGWI"
gene            complement(8345. .8533)
                /gene="SAV1323"
CDS             complement(8345. .8533)
                /gene="SAV1323"
                /codon_start=1
```

```
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAB57485.1"
                     /db_xref="GI:14247094"
                     /translation="MNISNENTVTLIAVIIAIIIGIFLQIFFKLPLIVTAVLSILLGI
                     FVGFIVYLIVSFYNKRKN"
     gene            8672. .9205
                     /gene="nuc"
     CDS             8672. .9205
                     /gene="nuc"
                     /note="SAV1324"
                     /codon_start=1
                     /transl_table=11
                     /product="thermonuclease"
                     /protein_id="BAB57486.1"
                     /db_xref="GI:14247095"
                     /translation="MKSNKSLAMIVVAIIIVGVLAFQFMNHTGPFKKGTNHETVQDLN
```

Alignment Scores:

Pred. No.:	1.56e-112	Length:	346900
Score:	1444.00	Matches:	285
Percent Similarity:	99.65%	Conservative:	1
Best Local Similarity:	99.30%	Mismatches:	1
Query Match:	99.38%	Indels:	0
DB:	1	Gaps:	0

US-09-925-637-64 (1-287) x AP003362 (1-346900)

```
Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251159 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 251100

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251099 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGCATGTTGTATTCATTAAATGCT 251040

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251039 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 250980

Qy         61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
              ||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db     250979 TATGAGTTAGGTGTGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 250920

Qy         81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250919 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCAT 250860

Qy        101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250859 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 250800

Qy        121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250799 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 250740

Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250739 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 250680

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250679 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 250620

Qy        181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250619 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAA 250560

Qy        201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250559 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 250500

Qy        221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250499 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 250440

Qy        241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250439 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 250380

Qy        261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250379 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 250320

Qy        281 LeuLeuGluIleValAspLeu 287
              |||||||||||||||||||||
Db     250319 TTATTAGAAATCGTTGATTTA 250299
```

```
RESULT 4
AR106466
LOCUS       AR106466                882 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 1 from patent US 6107058.
ACCESSION   AR106466
VERSION     AR106466.1  GI:12820996
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 882)
  AUTHORS   Gwynn,M. and Wilding,E.Imogen.
  TITLE     ispA from Staphylococcus aureus
  JOURNAL   Patent: US 6107058-A 1 22-AUG-2000;
FEATURES             Location/Qualifiers
     source          1. .882
                     /organism="unknown"
BASE COUNT      328 a    118 c    186 g    250 t
ORIGIN
```

Alignment Scores:

Pred. No.:	4.69e-115	Length:	882
Score:	1436.00	Matches:	284
Percent Similarity:	99.30%	Conservative:	1
Best Local Similarity:	98.95%	Mismatches:	2
Query Match:	98.83%	Indels:	0
DB:	6	Gaps:	0

US-09-925-637-64 (1-287) x AR106466 (1-882)

```
Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGCATGTTGTATTCATTAAATGCT 120

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy         61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
              ||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTGTGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy         81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
              ||||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAAGGAAATTAACAAATCAT 300

Qy        101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
```

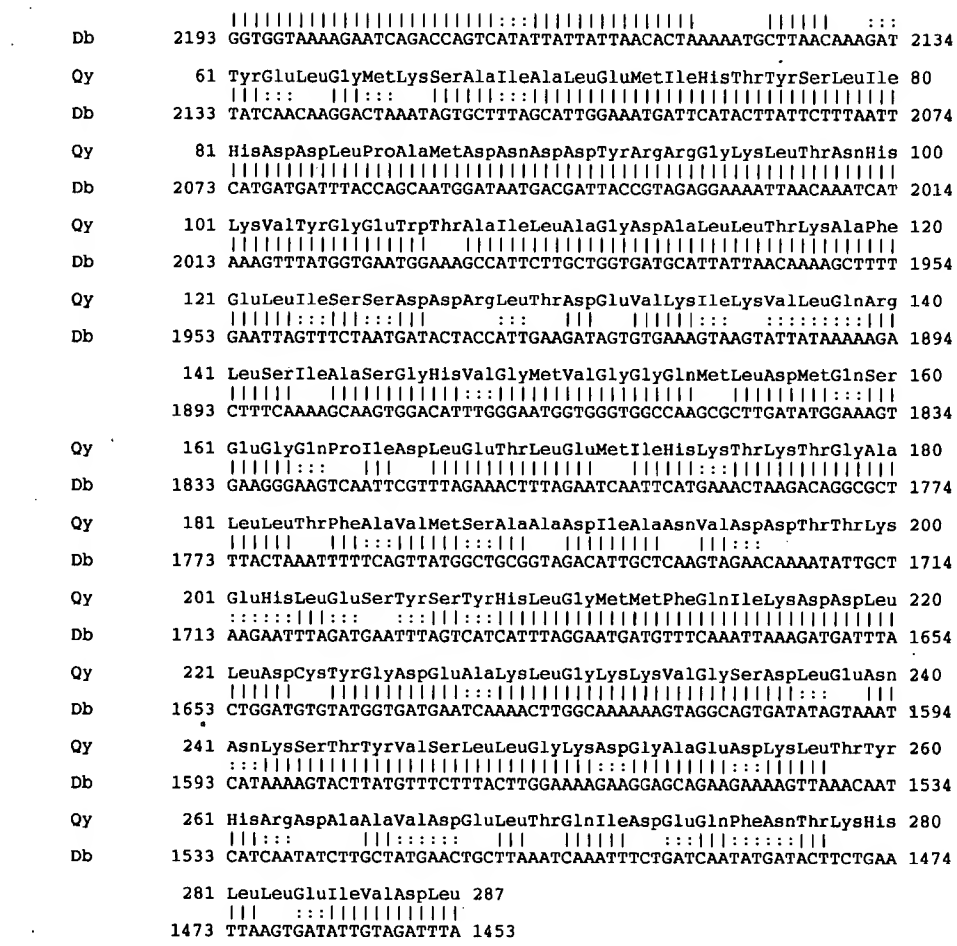
```
Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAA 600

Qy        201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660

Qy        221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Qy        241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy        261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy        281 LeuLeuGluIleValAspLeu 287
              |||||||||||||||||||||
Db        841 TTATTAGAAATCGTTGATTTA 861
```

```
RESULT 5
AF270104
LOCUS       AF270104               3444 bp    DNA     linear   BCT 01-AUG-2000
DEFINITION  Staphylococcus epidermidis strain SR1 clone step.1049b08 genomic
            sequence.
ACCESSION   AF270104
VERSION     AF270104.1  GI:9624010
KEYWORDS    .
SOURCE      Staphylococcus epidermidis.
  ORGANISM  Staphylococcus epidermidis
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1  (bases 1 to 3444)
  AUTHORS   Kimmerly,W.J., Taylor,J.David., Nelsen,A.J., Godlevski,M.M.,
            Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
            Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,
            Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
            Furdon,P.J.
  TITLE     Transposon-mediated sequencing of the Staphylococcus epidermidis
            genome
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3444)
  AUTHORS   Taylor,J.David., Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M.,
            Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
            Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,
            Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
            Furdon,P.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAY-2000) Departments of Genomic Sciences and
            Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
            Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES             Location/Qualifiers
     source          1. .3444
                     /organism="Staphylococcus epidermidis"
                     /strain="SR1"
                     /db_xref="taxon:1282"
                     /clone="step.1049b08"
BASE COUNT     1365 a    451 c    571 g   1057 t
ORIGIN

Alignment Scores:
Pred. No.:              4.62e-78        Length:         3444
```

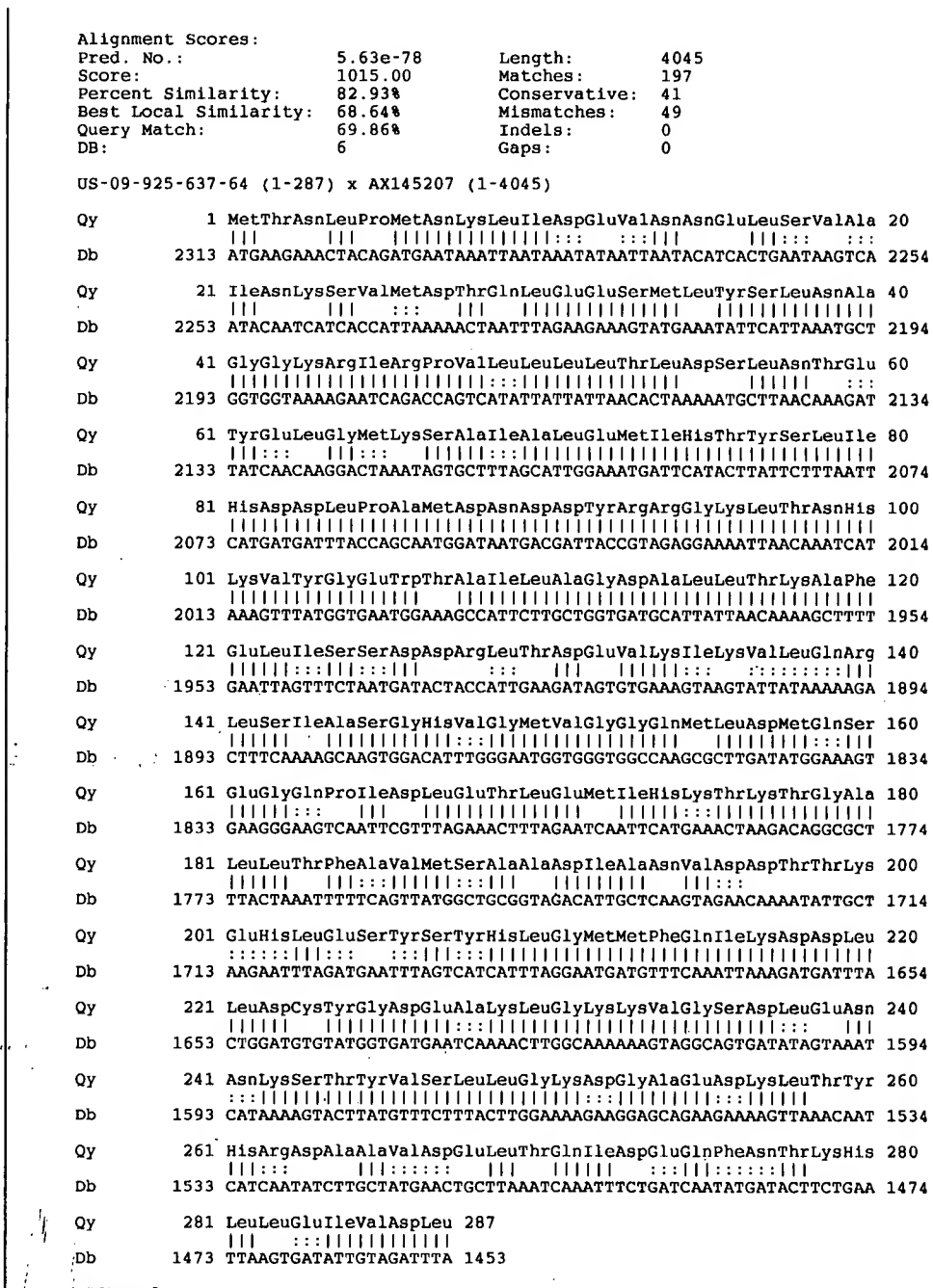
```
Score:                  1015.00         Matches:        197
Percent Similarity:     82.93%          Conservative:   41
Best Local Similarity:  68.64%          Mismatches:     49
Query Match:            69.86%          Indels:         0
DB:                     1               Gaps:           0

US-09-925-637-64 (1-287) x AF270104 (1-3444)

Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              |||       |||   |||||||||||||||:::   :::|||      |||:::   :::
Db        766 ATGAAGAAACTACAGATGAATAAATTAATAAATATAATTAATACATCACTGAATAAGTCA 825

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              |||      |||   :::   |||   ||||||||||||||   ||||||||||||||
Db        826 ATACAATCATCACCATTAAAAACTAATTTAGAAGAAAGTATGAAATATTCATTAAATGCT 885

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||:::|||||||||||||||      ||||||   :::
Db        886 GGTGGTAAAAGAATCAGACCAGTCATATTATTATTAACACTAAAAATGCTTAACAAAGAT 945

Qy         61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
              |||:::   |||:::   ||||||:::|||||||||||||||||||||||||||||||||
Db        946 TATCAACAAGGACTAAATAGTGCTTTAGCATTGGAAATGATTCATACTTATTCTTTAATT 1005

Qy         81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1006 CATGATGATTTACCAGCAATGGATAATGACGATTACCGTAGAGGAAAATTAACAAATCAT 1065

Qy        101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
              |||||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db       1066 AAAGTTTATGGTGAATGGAAAGCCATTCTTGCTGGTGATGCATTATTAACAAAAGCTTTT 1125

Qy        121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
              ||||||:::|||:::|||      :::   |||   |||||||:::   :::::::::|||
Db       1126 GAATTAGTTTCTAATGATACTACCATTGAAGATAGTGTGAAAGTAAGTATTATAAAAAGA 1185

Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
              ||||||   ||||||||||||:::||||||||||||||||||   |||||||||:::|||
Db       1186 CTTTCAAAAGCAAGTGGACATTTGGGAATGGTGGGTGGCCAAGCGCTTGATATGGAAAGT 1245

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
              ||||||:::   |||   |||||||||||||||   ||||||:::|||||||||||||||
Db       1246 GAAGGGAAGTCAATTCGTTTAGAAACTTTAGAATCAATTCATGAAACTAAGACAGGCGCT 1305

Qy        181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys 200
              ||||||   |||:::||||||:::|||   ||||||||   |||:::
Db       1306 TTACTAAATTTTTCAGTTATGGCTGCGGTAGACATTGCTCAAGTAGAACAAAATATTGCT 1365

Qy        201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
              ::::::|||:::   :::|||:::||||||||||||||||||||||||||||||||||||
Db       1366 AAGAATTTAGATGAATTTAGTCATCATTTAGGAATGATGTTTCAAATTAAAGATGATTTA 1425

Qy        221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
              ||||||   ||||||||||||:::|||||||||||||||||||||||||||:::   |||
Db       1426 CTGGATGTGTATGGTGATGAATCAAAACTTGGCAAAAAAGTAGGCAGTGATATAGTAAAT 1485

Qy        241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
              :::||||||||||||||||||||||||||||:::|||||||||:::||||||
Db       1486 CATAAAAGTACTTATGTTTCTTTACTTGGAAAAGAAGGAGCAGAAGAAAAGTTAAACAAT 1545

Qy        261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
              |||:::      |||::::::   |||   ||||||   :::|||:::::|||
Db       1546 CATCAATATCTTGCTATGAACTGCTTAAATCAAATTTCTGATCAATATGATACTTCTGAA 1605

Qy        281 LeuLeuGluIleValAspLeu 287
              |||   :::||||||||||||||
Db       1606 TTAAGTGATATTGTAGATTTA 1626
```

```
RESULT 6
AX145422
LOCUS       AX145422               3444 bp    DNA     linear   PAT 31-MAY-2001
DEFINITION  Sequence 4144 from Patent WO0134809.
ACCESSION   AX145422
```

```
VERSION     AX145422.1  GI:14283987
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 3444)
  AUTHORS   Kimmerly,W.J.
  TITLE     Staphylococcus epidermidis nucleic acids and proteins
  JOURNAL   Patent: WO 0134809-A 4144 17-MAY-2001;
            GLAXO GROUP LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .3444
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic nucleic acid sequence"
BASE COUNT     1365 a    451 c    571 g   1057 t
ORIGIN

Alignment Scores:
Pred. No.:              4.62e-78      Length:         3444
Score:                  1015.00       Matches:        197
Percent Similarity:     82.93%        Conservative:   41
Best Local Similarity:  68.64%        Mismatches:     49
Query Match:            69.86%        Indels:         0
DB:                     6             Gaps:           0

US-09-925-637-64 (1-287) x AX145422 (1-3444)

Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              |||      |||   |||||||||||||||:::   :::|||      |||:::   :::
Db        766 ATGAAGAAACTACAGATGAATAAATTAATAAATATAATTAATACATCACTGAATAAGTCA 825

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              |||      |||   :::   |||   ||||||||||||||   ||||||||||||||
Db        826 ATACAATCATCACCATTAAAAACTAATTTAGAAGAAAGTATGAAATATTCATTAAATGCT 885

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||:::|||||||||||||||      ||||||   :::
Db        886 GGTGGTAAAAGAATCAGACCAGTCATATTATTATTAACACTAAAAATGCTTAACAAAGAT 945

Qy         61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
              |||:::   |||:::   ||||||:::|||||||||||||||||||||||||||||||
Db        946 TATCAACAAGGACTAAATAGTGCTTTAGCATTGGAAATGATTCATACTTATTCTTTAATT 1005

Qy         81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1006 CATGATGATTTACCAGCAATGGATAATGACGATTACCGTAGAGGAAAATTAACAAATCAT 1065

Qy        101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
              ||||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db       1066 AAAGTTTATGGTGAATGGAAAGCCATTCTTGCTGGTGATGCATTATTAACAAAAGCTTTT 1125

Qy        121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
              ||||||:::|||:::|||      :::   |||   ||||||:::   ::::::::::|||
Db       1126 GAATTAGTTTCTAATGATACTACCATTGAAGATAGTGTGAAAGTAAGTATTATAAAAAGA 1185

Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
              ||||||   |||||||||||:::|||||||||||||||||   |||||||||:::|||
Db       1186 CTTTCAAAAGCAAGTGGACATTTGGGAATGGTGGGTGGCCAAGCGCTTGATATGGAAAGT 1245

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
              ||||||:::   |||   ||||||||||||||   ||||||:::||||||||||||||||
Db       1246 GAAGGGAAGTCAATTCGTTTAGAAACTTTAGAATCAATTCATGAAACTAAGACAGGCGCT 1305

Qy        181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys 200
              ||||||   |||:::||||||:::|||   ||||||||   |||:::
Db       1306 TTACTAAATTTTTCAGTTATGGCTGCGGTAGACATTGCTCAAGTAGAACAAAATATTGCT 1365

Qy        201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
              ::::::|||:::   :::|||:::||||||||||||||||||||||||||||||||||||
Db       1366 AAGAATTTAGATGAATTTAGTCATCATTTAGGAATGATGTTTCAAATTAAAGATGATTTA 1425

Qy        221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
              ||||||   ||||||||||||:::||||||||||||||||||||||||||||:::   |||
Db       1426 CTGGATGTGTATGGTGATGAATCAAAACTTGGCAAAAAAGTAGGCAGTGATATAGTAAAT 1485

Qy        241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
              :::||||||||||||||||||||||||||||||:::|||||||||:::||||||
Db       1486 CATAAAAGTACTTATGTTTCTTTACTTGGAAAAGAAGGAGCAGAAGAAAAGTTAAACAAT 1545

Qy        261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
              |||:::      |||::::::   |||   ||||||  .:::|||:::::|||
Db       1546 CATCAATATCTTGCTATGAACTGCTTAAATCAAATTTCTGATCAATATGATACTTCTGAA 1605

Qy        281 LeuLeuGluIleValAspLeu 287
              |||   :::|||||||||||
Db       1606 TTAAGTGATATTGTAGATTTA 1626

RESULT 7
AF269889/c
LOCUS       AF269889                4045 bp    DNA     linear   BCT 01-AUG-2000
DEFINITION  Staphylococcus epidermidis strain SR1 clone step.1028f08 genomic
            sequence.
ACCESSION   AF269889
VERSION     AF269889.1  GI:9623789
KEYWORDS    .
SOURCE      Staphylococcus epidermidis.
  ORGANISM  Staphylococcus epidermidis
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1  (bases 1 to 4045)
  AUTHORS   Kimmerly,W.J., Taylor,J.David., Nelsen,A.J., Godlevski,M.M.,
            Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
            Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,
            Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
            Furdon,P.J.
  TITLE     Transposon-mediated sequencing of the Staphylococcus epidermidis
            genome
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4045)
  AUTHORS   Taylor,J.David., Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M.,
            Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
            Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,
            Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
            Furdon,P.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAY-2000) Departments of Genomic Sciences and
            Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
            Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES             Location/Qualifiers
     source          1. .4045
                     /organism="Staphylococcus epidermidis"
                     /strain="SR1"
                     /db_xref="taxon:1282"
                     /clone="step.1028f08"
BASE COUNT     1249 a    691 c    536 g   1569 t
ORIGIN

Alignment Scores:
Pred. No.:              5.63e-78      Length:         4045
Score:                  1015.00       Matches:        197
Percent Similarity:     82.93%        Conservative:   41
Best Local Similarity:  68.64%        Mismatches:     49
Query Match:            69.86%        Indels:         0
DB:                     1             Gaps:           0

US-09-925-637-64 (1-287) x AF269889 (1-4045)

Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              |||      |||   |||||||||||||||:::   :::|||      |||:::   :::
Db       2313 ATGAAGAAACTACAGATGAATAAATTAATAAATATAATTAATACATCACTGAATAAGTCA 2254

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              |||      |||   :::   |||   ||||||||||||||   ||||||||||||||
Db       2253 ATACAATCATCACCATTAAAAACTAATTTAGAAGAAAGTATGAAATATTCATTAAATGCT 2194

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
```



RESULT 8
AX145207/c
LOCUS AX145207 4045 bp DNA linear PAT 31-MAY-2001
DEFINITION Sequence 3929 from Patent WO0134809.
ACCESSION AX145207
VERSION AX145207.1 GI:14283772
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 4045)
AUTHORS Kimmerly,W.J.
TITLE Staphylococcus epidermidis nucleic acids and proteins
JOURNAL Patent: WO 0134809-A 3929 17-MAY-2001;
GLAXO GROUP LIMITED (GB)
FEATURES Location/Qualifiers
source 1. .4045
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="synthetic nucleic acid sequence"
BASE COUNT 1249 a 691 c 536 g 1569 t
ORIGIN



Alignment Scores:

Pred. No.:	5.63e-78	Length:	4045
Score:	1015.00	Matches:	197
Percent Similarity:	82.93%	Conservative:	41
Best Local Similarity:	68.64%	Mismatches:	49
Query Match:	69.86%	Indels:	0
DB:	6	Gaps:	0

US-09-925-637-64 (1-287) x AX145207 (1-4045)

RESULT 9
AX141501

```
LOCUS       AX141501               783 bp    DNA     linear   PAT 31-MAY-2001
DEFINITION  Sequence 223 from Patent WO0134809.
ACCESSION   AX141501
VERSION     AX141501.1  GI:14281622
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 783)
  AUTHORS   Kimmerly,W.J.
  TITLE     Staphylococcus epidermidis nucleic acids and proteins
  JOURNAL   Patent: WO 0134809-A 223 17-MAY-2001;
            GLAXO GROUP LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .783
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic nucleic acid sequence"
BASE COUNT      297 a     93 c    147 g    246 t
ORIGIN
```

Alignment Scores:

Pred. No.:	1.17e-72	Length:	783
Score:	943.00	Matches:	181
Percent Similarity:	85.43%	Conservative:	36
Best Local Similarity:	71.26%	Mismatches:	37
Query Match:	64.90%	Indels:	0
DB:	6	Gaps:	0

US-09-925-637-64 (1-287) x AX141501 (1-783)

```
Qy         34 MetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgProValLeuLeuLeuLeuThr 53
              |||   |||||||||||||||||||||||||||||||||||||||:::|||||||||||||
Db          1 ATGAAATATTCATTAAATGCTGGTGGTAAAAGAATCAGACCAGTCATATTATTATTAACA 60

Qy         54 LeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMet 73
              |||      ||||||   :::|||:::   |||:::   ||||||:::|||||||||||||
Db         61 CTAAAAATGCTTAACAAAGATTATCAACAAGGACTAAATAGTGCTTTAGCATTGGAAATG 120

Qy         74 IleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArg 93
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATTCATACTTATTCTTTAATTCATGATGATTTACCAGCAATGGATAATGACGATTACCGT 180

Qy         94 ArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAsp 113
              ||||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db        181 AGAGGAAAATTAACAAATCATAAAGTTTATGGTGAATGGAAAGCCATTCTTGCTGGTGAT 240

Qy        114 AlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAspArgLeuThrAspGluVal 133
              |||||||||||||||||||||||||||:::|||:::|||      :::   |||   |||
Db        241 GCATTATTAACAAAAGCTTTTGAATTAGTTTCTAATGATACTACCATTGAAGATAGTGTG 300

Qy        134 LysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGly 153
              |||:::   :::::::::|||||||||   ||||||||||||:::||||||||||||||||
Db        301 AAAGTAAGTATTATAAAAAGACTTTCAAAAGCAAGTGGACATTTGGGAATGGTGGGTGGC 360

Qy        154 GlnMetLeuAspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIle 173
              |||   ||||||||:::|||||||||:::   |||   ||||||||||||||   |||
Db        361 CAAGCGCTTGATATGGAAAGTGAAGGGAAGTCAATTCGTTTAGAAACTTTAGAATCAATT 420

Qy        174 HisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAla 193
              |||:::||||||||||||||||||   |||:::||||||:::|||   |||||||||
Db        421 CATGAAACTAAGACAGGCGCTTTACTAAATTTTTCAGTTATGGCTGCGGTAGACATTGCT 480

Qy        194 AsnValAspAspThrThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMet 213
                 |||:::          ::::::|||:::   :::|||:::||||||||||||||||
Db        481 CAAGTAGAACAAAATATTGCTAAGAATTTAGATGAATTTAGTCATCATTTAGGAATGATG 540

Qy        214 PheGlnIleLysAspAspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLys 233
              ||||||||||||||||||||||||||   ||||||||||||:::||||||||||||||||
Db        541 TTTCAAATTAAAGATGATTTACTGGATGTGTATGGTGATGAATCAAAACTTGGCAAAAAA 600

Qy        234 ValGlySerAspLeuGluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGly 253
              ||||||||||||:::   |||:::|||||||||||||||||||||||||||||:::|||
```

```
Db        601 GTAGGCAGTGATATAGTAAATCATAAAAGTACTTATGTTTCTTTACTTGGAAAAGAAGGA 660

Qy        254 AlaGluAspLysLeuThrTyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAsp 273
              ||||||:::||||||      |||:::      |||:::::::   |||   ||||||
Db        661 GCAGAAGAAAAGTTAAACAATCATCAATATCTTGCTATGAACTGCTTAAATCAAATTTCT 720

Qy        274 GluGlnPheAsnThrLysHisLeuLeuGluIleValAspLeu 287
              :::|||:::::::|||      |||   :::||||||||||||
Db        721 GATCAATATGATACTTCTGAATTAAGTGATATTGTAGATTTA 762

RESULT 10
AL591978
LOCUS       AL591978               250050 bp    DNA     linear   BCT 06-JUN-2002
DEFINITION  Listeria monocytogenes strain EGD, complete genome, segment 6/12.
ACCESSION   AL591978 AL591824
VERSION     AL591978.1  GI:16410540
KEYWORDS    .
SOURCE      Listeria monocytogenes.
  ORGANISM  Listeria monocytogenes
            Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
REFERENCE   1
  AUTHORS   Glaser,P., Frangeul,L., Buchrieser,C., Rusniok,C., Amend,A.,
            Baquero,F., Berche,P., Bloecker,H., Brandt,P., Chakraborty,T.,
            Charbit,A., Chetouani,F., Couve,E., de Daruvar,A., Dehoux,P.,
            Domann,E., Dominguez-Bernal,G., Duchaud,E., Durant,L.,
            Dussurget,O., Entian,K.D., Fsihi,H., Portillo,F.G., Garrido,P.,
            Gautier,L., Goebel,W., Gomez-Lopez,N., Hain,T., Hauf,J.,
            Jackson,D., Jones,L.M., Kaerst,U., Kreft,J., Kuhn,M., Kunst,F.,
            Kurapkat,G., Madueno,E., Maitournam,A., Vicente,J.M., Ng,E.,
            Nedjari,H., Nordsiek,G., Novella,S., de Pablos,B., Perez-Diaz,J.C.,
            Purcell,R., Remmel,B., Rose,M., Schlueter,T., Simoes,N.,
            Tierrez,A., Vazquez-Boland,J.A., Voss,H., Wehland,J. and Cossart,P.
  TITLE     Comparative genomics of Listeria species
  JOURNAL   Science 294 (5543), 849-852 (2001)
  MEDLINE   21537279
   PUBMED   11679669
REFERENCE   2  (bases 1 to 250050)
  AUTHORS   Glaser,P., Frangeul,L. and Rusniok,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUN-2001) Glaser P., Institut Pasteur, Genomique des
            Microorganismes Pathogenes, 25 rue du Docteur Roux, 75724 Paris
            Cedex 15, FRANCE
COMMENT     E-mail: pglaser@pasteur.fr
            Phone: +33 1 45 68 89 96, Fax: +33 (0)1 45 68 87 86.
FEATURES             Location/Qualifiers
     source          1. .250050
                     /organism="Listeria monocytogenes"
                     /strain="EGD-e"
                     /db_xref="taxon:1639"
     gene            210. .753
                     /gene="lmo1125"
     RBS             210. .215
                     /gene="lmo1125"
     CDS             220. .753
                     /gene="lmo1125"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC99203.1"
                     /db_xref="GI:16410541"
                     /db_xref="SPTREMBL:Q8Y7Z4"
                     /translation="MLIPSYLKSTAKELSINDFLNVEIVTINKEETFEILYCGTLKEI
                     EGDQLITGEAYEIPLKIIAKSTLSGKEILLYDGAYYGYDSMFCDEFEEEATQNRELKK
                     YPINNLSNIRLSIGIGIDYESEKEEYEFDENGHVVLIDDRHIPWEQVKTDGFDFLEIT
                     ATDENGDSLLILTEELA"
     gene            join(764. .769,778. .1230)
                     /gene="lmo1126"
     RBS             764. .769
                     /gene="lmo1126"
     CDS             778. .1230
                     /gene="lmo1126"
                     /note="similar to E. coli YjaB protein"
                     /codon_start=1
                     /transl_table=11
```

```
                /protein_id="CAC99204.1"
                /db_xref="GI:16410542"
                /db_xref="SPTREMBL:Q8Y7Z3"
                /translation="MILINKRAENQDYQMILAIWEKSVIATHDFLAAEDRQFYKEQIP
                LFLDNVELFLWFAGEEIVGFSGTSERELDMLFLNPIATGNGYGSQILIWLIENKRINL
                VDVNEQNENATRFYLKHGFVVSSRSDLDGFGNPYPILHLQLKERVNGI"
gene            1208. .1213
                /gene="lmo1127"
RBS             1208. .1213
                /gene="lmo1127"
gene            1220. .1692
                /gene="lmo1127"
CDS             1220. .1675
                /gene="lmo1127"
                /codon_start=1
                /transl_table=11
                /protein_id="CAC99205.1"
                /db_xref="GI:16410543"
                /db_xref="SPTREMBL:Q8Y7Z2"
                /translation="MEYEQLLQNELGELEIIQEELQNKEYRGMIVKTAKDGTTIRTRL
                AKKTDKKAGYFVAFWEKDEFGKNCPFDENDCADFLCIVVIDGDLNGLFVFPRQCLIDK
                GILTSEASIGKMAARFYPTWCQDLNKTAKKTQAWQQVYFKNYSIRKEDH"
terminator      1660. .1692
                /gene="lmo1127"
gene            complement(1689. .2711)
                /gene="lmo1128"
CDS             complement(1689. .2711)
                /gene="lmo1128"
                /codon_start=1
                /transl_table=11
                /protein_id="CAC99206.1"
                /db_xref="GI:16410544"
                /db_xref="SPTREMBL:Q8Y7Z1"
                /translation="MAKNGVDTIKRWKKITIIMSAIIVILVAILFIGNNYEMNETRVA
                IPTTGGKLSAVVTTPKHGKPKGIIVFVHGDGAQEATQNGGYKPLMERFAKQGYISVSW
                DKLGVGKSSGNWLNQSMDDRANEVNQVIEWMKVKYPDSTAKIGLWGASQAGWVVPKAM
                NANNEIDFSILAAPAINWMRQGEYNTCWQAKDTGKTTKEVKQAQQNFLTDSALISKNK
                TYESYKQNGGKEEMTPGRYNFIRKNMNQDATLDLSKVKTKIYLALAEKDKNVDSAETK
                DVYTQQIKKENLETQTIPNVSHQMINPAIANSELLTNIVGLMVPKYFLVDNAYLDYCE
                QVVSKQ"
gene            complement(2666. .3280)
                /gene="lmo1129"
CDS             complement(2666. .3280)
                /gene="lmo1129"
                /note="similar to unknown proteins"
                /codon_start=1
                /transl_table=11
                /protein_id="CAC99207.1"
                /db_xref="GI:16410545"
                /db_xref="SPTREMBL:Q8Y7Z0"
                /translation="MTLSIKSIDSFEKDSIVEEVASLVFAGFKSKFTKTLFKETEAHD
                ITYAFCHYIYTEKGENLLIAIQDELVCGCLFVTSKSDSYQKLYQTLRKFLSFSQCLKL
                IFLLSLLSHKPVQNERYIDFITVSPNFRGQGIGKTLISHCLETFHDEKITLYVAKKNS
                GAYKLYRNLGFQVIQKENSTLMGVLTGIKEWRKMEWIQLNVGKK"
gene            complement(join(3277. .3996,4004. .4009))
                /gene="lmo1130"
CDS             complement(3277. .3996)
                /gene="lmo1130"
                /note="similar to transcription regulators"
                /codon_start=1
                /transl_table=11
                /protein_id="CAC99208.1"
                /db_xref="GI:16410546"
                /db_xref="SPTREMBL:Q8Y7Y9"
                /translation="MLTLSTGEIAQLFQISKYKIRHYIDEGILVPKRNPENGYYYFEE
                ADIYRLYQIILFRKIGFSIQEIKESLLGEKVTPMLEQAEQDLQQKIEELVEIQKTIQK
                IIHSQKEMRLNEITFVAKDDRYFKKVPNQIIDHNTINYSQAAKHNLPDLEEPFYIFSK
                QDTGVICLKSIKKRSDYTFPAGEYACKSFIAKDEATIETQIQLFLEELKYPTTSMILY
                ENIYPSLAYPDAMVYTMEVPL"
gene            complement(3288. .3293)
                /gene="lmo1129"
RBS             complement(3288. .3293)
                /gene="lmo1129"
RBS             complement(4004. .4009)
                /gene="lmo1130"
gene            join(4233. .4237,4245. .5960)
                /gene="lmo1131"
RBS             4233. .4237
                /gene="lmo1131"
CDS             4245. .5960
                /gene="lmo1131"
                /note="similar to ABC transporters, ATP-binding proteins"
                /codon_start=1
                /transl_table=11
                /protein_id="CAC99209.1"
                /db_xref="GI:16410547"
                /db_xref="SPTREMBL:Q8Y7Y8"
                /translation="MIDKRLFQLVEKKSLVLLILFRVLSLGLMIGLWLVFAQQLTHYL
                EGQNVDWLWLIGTVLVVLVGKAILTKLAEKQIYQASAELRLSMRRAVMEKAFRLGNNE
                GQLPASTLTQLAVDGIEQLEIYYSRFLPQLFYCLIASLMIFGSLVGFAWQPAIVLLIC
                MPMIPIVIMAVMKIAKRILSGYWSDYTNLGTKFHENLSGLSILKAYDQDKYKQEEIVS
                DAERFRKATMSLLSMQLNSITIMDIISYSGAALGIGMSLIMFTNGTISMTGMLMFLLL
                SAEFFIPMRQLGSLFHVAMNGISACSKLFAYLELKEQVYGSVELAKPLEKIEVHHLTY
                TYEGGKAKALQDVSVHFNKGSFSALVGKSGSGKSTFVRVLLNQLPGYQGEIVWNDVPL
                ASLSGEAIRKQAVLVDNHGYLYANSIRENLLIGNPAANDSDLWNVLEQVSLADFVRKL
                PEQLNENLEEDGSNLSGGQRQRLLLARALLRQAEVYVFDEITSGVDLESEKIILLVLQ
                ELAKEKIVLFISHRLYNVLGADQVLVFDAGKLLEVASPERLQQESNYFKNYFLEEEAL
                LKGGA"
gene            5948. .5953
                /gene="lmo1132"
RBS             5948. .5953
                /gene="lmo1132"
gene            5962. .7608
                /gene="lmo1132"
CDS             5962. .7608
                /gene="lmo1132"
                /note="similar to ABC transporters, ATP-binding proteins"
                /codon_start=1
                /transl_table=11
                /protein_id="CAC99210.1"
                /db_xref="GI:16410548"
                /db_xref="SPTREMBL:Q8Y7Y7"
                /translation="MSEWTIIGWLLKFVKPLRGKMILAILLGIISNLSVIMISLIGTY
                GIIAVILQQPLNPYKWLFVMVACGVVRGLARYLEQYLNHDIAFRLLAIIRERIFATLR
                KLGPARLSGKKSGDLVAAITTDVEALEVFFAHTISPVFIALGTTIATVGFLASYDVGL
                AIILLVGQIMVGVVLPMISYKRNKKIGTAYQQAFVGLNQMVMENIASLQDIFQFKLGE
                ERLRKLETQGEKLNRQYKKRIRQGSELQILGEWVLIGTATIILVLGSLSQLPLETVLI
                GTVLSLSSFGSVLALNALGTALLTTFASGKRLYALTEEKPVVLFNGQLELSDFESAEL
                DKVSFSHDGKQPLLNELSLALPKGKWLGIGGESGSGKSTLVKLLMRYWDPDGQVKLNN
                NELPKITESSLHRLEGVMEQRTFLFEDTLGNNIRLGKKDATLEEVKEAARKAAIDTWI
                ETLPEGYETIIGGQSRNLSDGERQRIGLARLFLHDAPLLLLDEPTSNLDYINEQAILN
                TLRSEIQDKTVLVISHRATTLDLAEEQLFIENGALKRAAK"
gene            complement(join(7604. .7637,7645. .7953,7961. .7966))
                /gene="lmo1133"
terminator      complement(7604. .7637)
                /gene="lmo1133"
gene            7615. .7637
                /gene="lmo1132"
terminator      7615. .7637
                /gene="lmo1132"
CDS             complement(7645. .7953)
                /gene="lmo1133"
                /note="similar to B. subtilis YjcS protein"
                /codon_start=1
                /transl_table=11
                /protein_id="CAC99211.1"
                /db_xref="GI:16410549"
                /db_xref="SPTREMBL:Q8Y7Y6"
                /translation="MYLKENTNYLYCTAVIQTTGKVSYEQLVEHLEALRAKTVKEPGC
```

Alignment Scores:

Pred. No.:	7.64e-49	Length:	250050
Score:	702.50	Matches:	148
Percent Similarity:	69.29%	Conservative:	46
Best Local Similarity:	52.86%	Mismatches:	81
Query Match:	48.35%	Indels:	5

```
DB:                 1          Gaps:         2

US-09-925-637-64 (1-287) x AL591978 (1-250050)

Qy          8 LysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerValMetAsp 27
              |||:::|||||||||           |||         ||||||:::      ::::::
Db     225970 AAAGTGATTGATGAG---------TCGCTTTTTAAAGAAATAAACGAGCGAAATATCGAA 226020

Qy         28 ThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgPro 47
              :::|||:::|||||||||||||||:::   ||||||||||||||||||||||
Db     226021 CCTAGACTAAAAGAGTCCATGTTATATTCTATTCAAGCAGGCGGGAAACGGATTCGTCCA 226080

Qy         48 ValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSer 67
              :::|||:::      ||||||  :::|||                ||||||:::|||:::
Db     226081 ATGCTAGTTTTCGCTACGCTTCAAGCTTTAAAAGTAAATCCGCTTTTAGGTGTAAAAACC 226140

Qy         68 AlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMet 87
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     226141 GCAACAGCGTTAGAAATGATTCACACATACAGCCTTATTCATGATGATCTACCAGCAATG 226200

Qy         88 AspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrpThr 107
              ||||||||||||||||||||||||||  ||||||||||||||||:::|||:::   |||
Db     226201 GATAACGATGATTATCGTCGAGGCAAGTGGACTAATCATAAAGTTTTTGGCGATGCAACT 226260

Qy        108 AlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAsp 127
              |||||||||||||||||||||||||||||  ||||||  :::::::::  ||||||
Db     226261 GCGATTTTGGCAGGAGATGCTTTACTAACGCTCGCTTTTTCTATTTTAGCTGAAGACGAT 226320

Qy        128 ArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHis 147
              |||:::   |||   :::|||   ::::::   ::::::|||   :::||||||
Db     226321 AATTTATCTTTTGAGACACGCATTGCTTTGATTAACCAAATTAGTTTTAGTAGCGGTGCA 226380

Qy        148 ValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIleAspLeu 167
              ||||||||||||||||||:::   |||:::::::::|||   :::   :::   |||
Db     226381 GAAGGAATGGTTGGTGGTCAACTTGCAGACTTGGAAGCGGAAAACAAACAAGTGACGCTA 226440

Qy        168 GluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMet 187
              |||   |||      ||||||     ||||||||   ||||||   :::||||||
Db     226441 GAAGAGTTATCATCCATTCATGCACGAAAAACGGGTGAATTATTAATTTATGCTGTAACC 226500

Qy        188 SerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyrSer 207
              |||||||||   ||||||      |||           :::   |||      ::::::
Db     226501 TCTGCAGCAAAAATTGCGGAAGCTGATCCAGAACAAACGAAACGCTTACGAATTTTTGCA 226560

Qy        208 TyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGlyAspGlu 227
              ::::::|||:::   |||||||||   ||||||:::||||||      |||||||||
Db     226561 GAGAATATTGGGATTGGATTTCAAATTAGCGACGATATTTTAGATGTAATTGGTGATGAA 226620

Qy        228 AlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyrValSer 247
              |||:::|||||||||   |||:::|||         ||||||||||||||
Db     226621 ACGAAAATGGGTAAAAAGACAGGGGCCGACGCTTTTCTGAATAAAAGTACCTATCCCGGA 226680

Qy        248 LeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAlaValAsp 267
              ||||||      |||||||||:::      |||       |||          |||
Db     226681 TTACTCACGCTTGATGGGGCAAAAAGGGCATTAAATGAGCATGTTACGATTGCAAAGTCA 226740

Qy        268 GluLeuThrGlnIleAspGluGlnPheAsnThrLysHisLeuLeuGluIleValAspLeu 287
              |||:::      |||      |||:::   :::   ||||||::::::   ||||||
Db     226741 GCGCTTTCAGGGCATGAT------TTCGATGATGAAATTCTCTTGAAACTTGCTGATTTA 226794

RESULT 11
AX416832
LOCUS       AX416832               5635 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 3823 from Patent WO0228891.
ACCESSION   AX416832
VERSION     AX416832.1  GI:21449341
KEYWORDS    .
SOURCE      Listeria monocytogenes ATCC 19115.
  ORGANISM  Listeria monocytogenes ATCC 19115
            Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
REFERENCE   1
  AUTHORS   Glaser,P. and Kunst,F.
  TITLE     Listeria innocua, genome and applications
  JOURNAL   Patent: WO 0228891-A 3823 11-APR-2002;
            Pasteur Institut (FR)
FEATURES             Location/Qualifiers
     source          1. .5635
                     /organism="Listeria monocytogenes ATCC 19115"
                     /db_xref="taxon:176281"
BASE COUNT     1896 a    927 c   1238 g   1574 t
ORIGIN

Alignment Scores:
Pred. No.:                2.8e-50        Length:          5635
Score:                    695.50         Matches:         148
Percent Similarity:       68.93%         Conservative:    45
Best Local Similarity:    52.86%         Mismatches:      82
Query Match:              47.87%         Indels:          5
DB:                       6              Gaps:            2

US-09-925-637-64 (1-287) x AX416832 (1-5635)

Qy          8 LysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerValMetAsp 27
              |||:::|||||||||           |||         |||:::::::      ::::::
Db       2914 AAAGTGATTGATGAG---------TCGCTTTTTAAAGAAATAAGCGAGCGAAATATCGAA 2964

Qy         28 ThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgPro 47
              :::|||:::|||||||||||||||:::   ||||||||||||||||||||||
Db       2965 CCTAGACTAAAAGAGTCCATGTTATATTCTGTTCAAGCAGGTGGAAAACGGATTCGTCCC 3024

Qy         48 ValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSer 67
              :::|||:::      ||||||  :::|||                ||||||:::|||:::
Db       3025 ATGCTAGTTTTTGCTACACTTCAAGCCTTAAAAGTAAATCCGCTTTTAGGTGTGAAAACA 3084

Qy         68 AlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMet 87
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3085 GCAACGGCGTTAGAAATGATTCATACGTATAGCCTGATCCATGATGATTTACCAGCAATG 3144

Qy         88 AspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrpThr 107
              ||||||||||||||||||||||||||  ||||||||||||||||:::|||:::   |||
Db       3145 GATAATGATGACTATCGTCGCGGCAAATATACTAATCATAAAGTTTTTGGGGATGCAACT 3204

Qy        108 AlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAsp 127
              |||||||||||||||||||||||||||||  ||||||  :::::::::  ||||||
Db       3205 GCGATTTTGGCAGGAGACGCTTTACTAACGCTTGCTTTTTCTATTTTAGCGGAAGACGAA 3264

Qy        128 ArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHis 147
              |||:::   |||   :::|||   ::::::   ::::::|||   :::||||||
Db       3265 AATTTATCTTTTGAAACACGCATAGCTTTGATTAATCAAATTAGTTTTAGTAGTGGTGCA 3324

Qy        148 ValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIleAspLeu 167
              ||||||||||||||||||:::   ||||||:::::::||   :::   :::   |||
Db       3325 GAAGGAATGGTTGGCGGTCAACTTGCAGATATGGAGGCAGAAAACAAACAAGTGACATTA 3384

Qy        168 GluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMet 187
              |||   |||      ||||||     ||||||||   ||||||   |||||||||
Db       3385 GAAGAACTATCATCCATCCATGCACGGAAAACAGGCGAATTATTAATTTTTGCCGTAACT 3444

Qy        188 SerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyrSer 207
              |||||||||   ||||||      |||           :::   |||      ::::::
Db       3445 TCTGCTGCGAAAATTGCAGAAGCTGACCCGGAACAAACGAAAAGATTACGGATTTTTGCG 3504

Qy        208 TyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGlyAspGlu 227
              ::::::|||:::   |||||||||   ||||||:::||||||      |||||||||
Db       3505 GAAAATATTGGGATTGGATTCCAAATCAGCGACGATATTTTAGATGTTATTGGTGATGAA 3564

Qy        228 AlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyrValSer 247
              |||:::|||||||||   |||   |||         ||||||||||||||
Db       3565 ACAAAAATGGGTAAAAAGACAGGGGTCGACGCTTTTCTGAACAAAAGTACCTATCCCGGA 3624

Qy        248 LeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAlaValAsp 267
              ||||||      |||||||||:::      |||       |||          |||
Db       3625 TTACTCACGCTTGATGGTGCCAAAAGGGCATTAAATGAGCATGTTGCGATTGCAAAGTCA 3684

Qy        268 GluLeuThrGlnIleAspGluGlnPheAsnThrLysHisLeuLeuGluIleValAspLeu 287
              |||:::      |||      |||:::   :::   ||||||::::::   ||||||
```

```
Db       3685 GCGCTTTCAGGGCATGAT------TTTGATGATGAAATTCTCTTGAAACTTGCTGATTTA 3738

RESULT 12
AX417035/c
LOCUS       AX417035              195269 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 4026 from Patent WO0228891.
ACCESSION   AX417035
VERSION     AX417035.1  GI:21449645
KEYWORDS    .
SOURCE      Listeria innocua.
  ORGANISM  Listeria innocua
            Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
REFERENCE   1
  AUTHORS   Glaser,P. and Kunst,F.
  TITLE     Listeria innocua, genome and applications
  JOURNAL   Patent: WO-0228891-A 4026 11-APR-2002;
            Pasteur Institut (FR)
FEATURES             Location/Qualifiers
     source          1. .195269
                     /organism="Listeria innocua"
                     /db_xref="taxon:1642"
                     /note="second part of seq 8-first part = seq 8 (000001 -
                     349980)-second part = seq 4026 (300001 - 495269)"
BASE COUNT    57412 a  39214 c  31781 g  66862 t
ORIGIN

Alignment Scores:
Pred. No.:              1.34e-47        Length:         195269
Score:                  686.50          Matches:        146
Percent Similarity:     68.21%          Conservative:   45
Best Local Similarity:  52.14%          Mismatches:     84
Query Match:            47.25%          Indels:         5
DB:                     6               Gaps:           2

US-09-925-637-64 (1-287) x AX417035 (1-195269)

Qy          8 LysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerValMetAsp 27
              |||:::::::||||||       |||       ||||||       ::::::
Db      66194 AAAGTACTTGATGAG---------TCGCTTTTTAAAGAAATAAATATGCGCAATATCGAA 66144

Qy         28 ThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgPro 47
                ::::|||:::|||||||||||||||||||:::   ||||||||||||||||||||||
Db      66143 CCTAAGTTGAAAGAATCAATGTTATATTCAGTGCAAGCTGGCGGAAAACGAATTCGTCCA 66084

Qy         48 ValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSer 67
              :::|||:::      ||||||   :::|||||||   |||         |||:::|||:::
Db      66083 ATGCTTGTTTTTGCAACGCTTCAAGCCTTAAATATTGAGCCAATGCGCGGTTTAAAAACA 66024

           68 AlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMet 87
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66023 GCTACGGCACTTGAAATGATTCATACGTACAGCTTAATTCACGATGATTTACCAGCAATG 65964

Qy         88 AspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrpThr 107
              |||:::|||||||||||||||||||||   ||||||||||||:::||||||:::   |||
Db      65963 GATGATGATGACTATCGACGCGGGAAATGGACGAACCATAAAATATACGGTGATGCAACA 65904

Qy        108 AlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAsp 127
              ||||||||||||||||||||||||||||   ||||||   :::::::::   |||:::
Db      65903 GCAATTTTAGCAGGAGATGCTTTGTTAACACTCGCTTTTTCTATTTTAGCTGAAGATGAA 65844

Qy        128 ArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHis 147
              |||:::   |||   :::|||   ::::::   :::::::|||   :::||||||
Db      65843 AATCTATCTTTCGAAACGCGTATTGCTTTAATTAATCAAATTAGTTATAGCAGTGGAGCA 65784

Qy        148 ValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIleAspLeu 167
                ||||||||||||||||||      ||||||:::::|||   :::   :::   |||
Db      65783 GAAGGCATGGTAGGTGGGCAACAAGCCGATATGGAAGCTGAAAATAAACAAGTCACATTA 65724

Qy        168 GluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMet 187
              |||   |||       ||||||       |||||||||   ||||||   |||||||||
Db      65723 GAAGAACTAGCATCAATCCACGCTCGCAAAACTGGTGAACTATTAATTTTCGCAGTAACC 65664

Qy        188 SerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyrSer 207
              |||||||||   ||||||                    :::   |||      ::::::
Db      65663 TCAGCCGCAAAAATCGCTGAAGCAACTCCAGAACAAACAAAACGATTACGAATTTTGCA 65604

Qy        208 TyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGlyAspGlu 227
                 :::::::||:::   |||||||||   ||||||:::||||||       |||||||||
Db      65603 GAAAATATCGGCATTGGTTTTCAAATTAGCGACGATATTTTAGATGTAATTGGCGATGAA 65544

Qy        228 AlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyrValSer 247
                 |||:::||||||||   |||   |||         ||||||||||||||
Db      65543 ACAAAAATGGGTAAAAAGACAGGGGTCGACGCTTTTCTGAATAAAAGTACCTATCCCGGA 65484

Qy        248 LeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAlaValAsp 267
              ||||||      :::||||||:::      |||      |||      |||
Db      65483 TTACTCACACTCGAAGGTGCTAAACGGGCATTAAATGAGCATGTTTCAATTGCAAAGTCA 65424

Qy        268 GluLeuThrGlnIleAspGluGlnPheAsnThrLysHisLeuLeuGluIleValAspLeu 287
                 |||:::      |||      |||:::   :::   ||||||:::::   ||||||
Db      65423 GCGCTTTCAGGACATGAT------TTCGACGATGAAATTCTCCTTAAACTTGCTGATTTA 65370

RESULT 13
AL596168
LOCUS       AL596168              333050 bp    DNA     linear   BCT 06-JUN-2002
DEFINITION  Listeria innocua Clip11262 complete genome, segment 6/12.
ACCESSION   AL596168 AL592022
VERSION     AL596168.1  GI:16413677
KEYWORDS    .
SOURCE      Listeria innocua.
  ORGANISM  Listeria innocua
            Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
REFERENCE   1
  AUTHORS   Glaser,P., Frangeul,L., Buchrieser,C., Rusniok,C., Amend,A.,
            Baquero,F., Berche,P., Bloecker,H., Brandt,P., Chakraborty,T.,
            Charbit,A., Chetouani,F., Couve,E., de Daruvar,A., Dehoux,P.,
            Domann,E., Dominguez-Bernal,G., Duchaud,E., Durant,L.,
            Dussurget,O., Entian,K.D., Fsihi,H., Portillo,F.G., Garrido,P.,
            Gautier,L., Goebel,W., Gomez-Lopez,N., Hain,T., Hauf,J.,
            Jackson,D., Jones,L.M., Kaerst,U., Kreft,J., Kuhn,M., Kunst,F.,
            Kurapkat,G., Madueno,E., Maitournam,A., Vicente,J.M., Ng,E.,
            Nedjari,H., Nordsiek,G., Novella,S., de Pablos,B., Perez-Diaz,J.C.,
            Purcell,R., Remmel,B., Rose,M., Schlueter,T., Simoes,N.,
            Tierrez,A., Vazquez-Boland,J.A., Voss,H., Wehland,J. and Cossart,P.
  TITLE     Comparative genomics of Listeria species
  JOURNAL   Science 294 (5543), 849-852 (2001)
  MEDLINE   21537279
   PUBMED   11679669
REFERENCE   2  (bases 1 to 333050)
  AUTHORS   Glaser,P., Frangeul,L. and Rusniok,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-2001) Glaser P., Institut Pasteur, Genomique des
            Microorganismes Pathogenes, 25 rue du Docteur Roux, 75724 Paris
            Cedex 15, FRANCE
COMMENT     E-mail: pglaser@pasteur.fr
            Phone: +33 (0)1 45 68 89 96, Fax: +33 (0)1 45 68 87 86.
FEATURES             Location/Qualifiers
     source          1. .333050
                     /organism="Listeria innocua"
                     /strain="Clip11262"
                     /db_xref="taxon:1642"
     gene            915. .4679
                     /gene="lin1204"
     RBS             915. .920
                     /gene="lin1204"
     CDS             927. .4640
                     /gene="lin1204"
                     /note="similar to internalin proteins (LPXTG motif)"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC96435.1"
                     /db_xref="GI:16413678"
                     /db_xref="SPTREMBL:Q92CG8"
                     /translation="MKRKISSIIVVGIMLFQSLTMYPFITEAKENEQKEEINKPSKIT
                     KGLTNSLKYTKTILETGDTYDSVFPDSALAKVVAKEATGSENTTQLVTQADLNKIKSL
                     NGYNKGISVLTGIDLLVNVTSISLNNNQVTDISPIDQLPNLVSLSVKNNQISSLILNA
```

```
                QNQLPKLTTIDIENNPDLNTIDIQDQPQLVDVKTSGYTGLRKLTTVIAKNNPELVNLG
                QYTIRNVYFSQVASLTKVELVNLPKVRKVNLERNSINELKVTDLAIEDLPLEENELTD
                TVFDNIQNLPNLKTLDLSKNQLEEVVLDKTDVENLPNLMTLDIQQNLAIKLINVQDQP
                QLVDVKTSDYKELSALTTVIAKNNPELVNLGQYTIRNVYFSQVASLTKVELVNLPKVR
                KVNLERNSINELKVTDLAIEDLPLEENELTDTVFDNIQNLPNLKTLDLSKNQLEEVVL
                DKTDVENLPNLMTLDIQQNLAIKLINVQDQPQLVDVKTSDYKELSALTTVIAKNNPEL
                VNLGYPIMQNVYFSQVASLTKVELANLPKVRAVRLERNSINQIELNNLVSVKDVNLNT
                NKITNDSIEKFRGMPILATLNLNKNQITNINMLDGFPQMTTLNIDLNSISVLPSNLKT
                KMPKLSRISALNQTVTLGKAIVVDDSDLIINNEISNFEKLTEPSPISNFGTYANEKIT
                WSSERIKNLTEVSFQFSELINVTGIDGTFSGKVTQPFKRSTTPVINADSEIHYPQGTE
                KTEAEFLKDVQAQTTDDLAIKSDFETMVNLKKVGKYTVTLNVENLDGIKATPKEVIVY
                IDAVQGANITVKYEDESGNKLAENSILTGNVGEEYSSSAKEILGYTLTEMPSNAQGEF
                SLEEQTVTYIYSKNPVPAKDITVQYTDEDGMELAPTETLSGNVDENYVTTAKTFTGYE
                LIETPSNAEGKFSENAQTVTYVYRAIKAEPILAKEVTVNYQDELGAKISETEVLTGEI
                GETYTTVAKTIDGYTLIKSPTNANGIFNENPQTVTYVYQLQNSPIAANITVKHLDENN
                NELAPSEVLSGNVDEAYTTKPKEIKDYSLVKVPTNASGKFTTEAQTVIYHYKKNSIPT
                SSYITVKYVDETGKELAISKVLNGNINDSYATTAKEIKGYTLVQKPANATGKFTDQNQ
                TIKYVYRANKAEVNPNPNSDVLGSNSNGDIQPTNDIKPTNDVQPTNEAPQTLPKTGSQ
                PANLIFGLGILLVLLSTQWLHRNKRKKAKSHNC"
terminator      4659. .4679
                /gene="lin1204"
gene            join(4883. .6241,6249. .6276)
                /gene="lin1205"
CDS             4883. .6241
                /gene="lin1205"
                /codon_start=1
                /transl_table=11
                /protein_id="CAC96436.1"
                /db_xref="GI:16413679"
                /db_xref="SPTREMBL:Q92CG7"
                /translation="MKKMLPESKVESIRKEGFLNRAVEAYRFFYPTVSNVSNFKALHD
                LGITENHDFIIQLTTPDLNVLTQNSDTPYCLGTGNTENGPVVIELPLGAIVGVADDIN
                FKFITNMGLTGDEEGKGAKYLYLPPNYDGDIPEGYIVRKPSSFRFLICLRAMVHQASD
                YEKAFELLKTVKFYPLKEAKNNPTSTFHDFSHRKAISTPYYVEGKFDYWEVIKWALDN
                DETDPEYYQMYGLLKAIGLAPNKEFNPDSDTKALLIEAAEKADKMMFVNSFNTDDPAA
                IVWPGKNWEWAVYGENNDFYEKTYLNLPVRERWFYQATLETHMMFMHKVGFGSVYMLG
                VKDKEGNYLDGGKSYTLKVPTPVPTSIFWSVTVYEMDTRSEIVTEQFMPALNSIKDTF
                EADADGNTTLYFGPNPPEDESLPWIQTVPDANWFTYFRIYGPTEPAFDNSWQLYDFEE
                VK"
gene            complement(join(6249. .6276,6283. .6876,6885. .6890))
                /gene="lin1206"
terminator      6249. .6276
                /gene="lin1205"
terminator      complement(6249. .6276)
                /gene="lin1206"
CDS             complement(6283. .6876)
                /gene="lin1206"
                /note="similar to B. subtilis YdeI protein"
                /codon_start=1
                /transl_table=11
                /protein_id="CAC96437.1"
                /db_xref="GI:16413680"
                /db_xref="SPTREMBL:Q92CG6"
                /translation="MAKTELNPKVDSFLNKPSTWQAEFKALREMAITFELEEEFKWGK
                PCYAINGSNVFLIHGFKNYCALLFMKGALLRDPDNILVQQTENVQAARQIRFTNLQEI
                LDQKEILKKYIQNAIEVEKAGLEVALKPRAETPIPEELLVKFEELPALKTAFEALTPG
                RQKAYLLYFAAPKQSKTRVSRIEKYEATILDGLGLND"
RBS             complement(6885. .6890)
                /gene="lin1206"
gene            complement(6994. .7414)
                /gene="lin1207"
CDS             complement(6994. .7401)
                /gene="lin1207"
                /codon_start=1
                /transl_table=11
                /protein_id="CAC96438.1"
                /db_xref="GI:16413681"
                /db_xref="SPTREMBL:Q92CG5"
                /translation="MTLNVYLNFRTQSRDAIAFYEEIFGTKCTDLMTYGEIESAEEPI
                EDSLKDLVMNASLVLDGVKVMFSDVPKSMPLTFGDNITLVIDTTDEIKLTKQFHQLAE
                GGNIIMPLGKTFWSEKYGQVTDKFGIGWQLNLS"
RBS             complement(7409. .7414)
                /gene="lin1207"
gene            7553. .8165
                /gene="lin1208"
RBS             7553. .7558
                /gene="lin1208"
CDS             7566. .8165
                /gene="lin1208"
                /note="weakly similar to phosphoglycerate mutase 1"
                /codon_start=1
                /transl_table=11
                /protein_id="CAC96439.1"
                /db_xref="GI:16413682"
                /db_xref="SPTREMBL:Q92CG4"
                /translation="MKKTLYLMRHGQTLFNQRKKIQGFCDAPLTDLGIKQAKIAGSYF
                KENNITFDQVYSSTSERACDTLELITDKSYQRLKGLKEWNFGAFEGESEDLNPPLPYG
                DFFAEYGGEREVDFRDRLVTTMERIMSQENHDTVLAVSHGAACAQFARYWEKTSKIGK
                VTGLKNCCILKFEYENGRFTLVNFINHDFENGTHMESVK"
gene            join(8347. .8352,8361. .8723)
                /gene="lin1209"
RBS             8347. .8352
                /gene="lin1209"
CDS             8361. .8723
                /gene="lin1209"
                /codon_start=1
                /transl_table=11
                /protein_id="CAC96440.1"
                /db_xref="GI:16413683"
                /db_xref="SPTREMBL:Q92CG3"
                /translation="MSKYDAEIVNAQQSITGQQEKIRRLKELITKIERKGEQISSIPI
                RQIDLPFDFWQGKSDSVMREVLQRRLQFWNNQNAVVNDLVQELRAAMTKMQNQISDYQ
                ADISYYTNLKQMEEEVNV"
gene            8701. .8706
                /gene="lin1210"
RBS             8701. .8706
                /gene="lin1210"
gene            8716. .8991
                /gene="lin1210"
CDS             8716. .8991
                /gene="lin1210"
                /codon_start=1
                /transl_table=11
                /protein_id="CAC96441.1"
                /db_xref="GI:16413684"
                /db_xref="SPTREMBL:Q92CG2"
                /translation="MSDIQVNESQIIKSMNQFAGKQKVLYKKSSSTSHNNKLATYHKI
                QTILTKMDTQMNQLQQLGERTTQDIEKNCAEFVNLDKNLLTDIEVTG"
gene            8992. .10532
                /gene="lin1211"
RBS             8992. .8997
                /gene="lin1211"
CDS             8999. .10459
                /gene="lin1211"
                /codon_start=1
                /transl_table=11
                /protein_id="CAC96442.1"
                /db_xref="GI:16413685"
                /db_xref="SPTREMBL:Q92CG1"
                /translation="MKIDVRELNEIVSENIVMLKKEYENLNIAMNATTNFTVETQEFF
                KGKTADASRAYLQEAYKPVQQKALEVNKLMTKILSEYIADAEAQFGANGMIDIQSIEM
                EYKRSFNQVSDDEMREYRDLNNLIQEANEFTYFQSSNLGFFEELHHDALTEITNIRMK
                LEDFETKWNVELSKAEKLQKELDWMLTQVNDNKIIPTSYQAGTLPFGKPFKSRDEFDT
                VQEYYQYVVEWAVANGKAKEIIVDGQRYYEFTDTFYYENYNKEVNGTLVMTEVNGEMI
                LFVQRKDIESAVSVPFGVFYPDPFQRLANGGETDGTIKLYHWKSKDIPEGLSVPNEEL
                KAMADFFADTQNSYDAGMAILKYLAHNKNPEKLSFTDTELDKVATNWSGGKNRFDFAG
                EVEAHADMLTDSILTELAAKPMDKIDRPGNTDEEKEQGYYHAASIADMQKGEDGMMNK
                SFNSIWGSGQYSAYLQNLMHGQENTPKAKKRSFPRQ"
terminator      10513. .10532
                /gene="lin1211"
gene            join(10896. .11450,11456. .11489)
                /gene="lin1212"
CDS             10896. .11450
                /gene="lin1212"
                /codon_start=1
                /transl_table=11
```

```
                     /protein_id="CAC96443.1"
                     /db_xref="GI:16413686"

Alignment Scores:
Pred. No.:                2.6e-47        Length:          333050
Score:                    686.50         Matches:         146
Percent Similarity:       68.21%         Conservative:    45
Best Local Similarity:    52.14%         Mismatches:      84
Query Match:              47.25%         Indels:          5
DB:                       1              Gaps:            2

US-09-925-637-64 (1-287) x AL596168 (1-333050)

Qy          8 LysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerValMetAsp 27
              |||:::::::||||||            |||         ||||||         ::::::
Db     174757 AAAGTACTTGATGAG---------TCGCTTTTTAAAGAAATAAATATGCGCAATATCGAA 174807

Qy         28 ThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgPro 47
              :::|||:::||||||||||||||||||:::   ||||||||||||||||||||||||||
       174808 CCTAAGTTGAAAGAATCAATGTTATATTCAGTGCAAGCTGGCGGAAAACGAATTCGTCCA 174867

             48 ValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSer 67
              :::|||:::      ||||||   :::||||||  |||         |||:::|||:::
Db     174868 ATGCTTGTTTTTGCAACGCTTCAAGCCTTAAATATTGAGCCAATGCGCGGTTTAAAAACA 174927

Qy         68 AlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMet 87
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     174928 GCTACGGCACTTGAAATGATTCATACGTACAGCTTAATTCACGATGATTTACCAGCAATG 174987

Qy         88 AspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrpThr 107
              |||:::||||||||||||||||||   ||||||||||:::|||||::   |||
Db     174988 GATGATGATGACTATCGACGCGGGAAATGGACGAACCATAAAATATACGGTGATGCAACA 175047

Qy        108 AlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAsp 127
              |||||||||||||||||||||||||||||   ||||||   :::::::::   |||:::
Db     175048 GCAATTTTAGCAGGAGATGCTTTGTTAACACTCGCTTTTTCTATTTTAGCTGAAGATGAA 175107

Qy        128 ArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHis 147
              |||:::   |||   :::|||   ::::::   :::::::|||   :::||||||
Db     175108 AATCTATCTTTCGAAACGCGTATTGCTTTAATTAATCAAATTAGTTATAGCAGTGGAGCA 175167

Qy        148 ValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIleAspLeu 167
              ||||||||||||||||||||      ||||||::::::|||   :::   :::   |||
Db     175168 GAAGGCATGGTAGGTGGGCAACAAGCCGATATGGAAGCTGAAAATAAACAAGTCACATTA 175227

Qy        168 GluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMet 187
              |||   |||      ||||||      |||||||||   ||||||   |||||||||
Db     175228 GAAGAACTAGCATCAATCCACGCTCGCAAAACTGGTGAACTATTAATTTTCGCAGTAACC 175287

Qy        188 SerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyrSer 207
              |||||||||   ||||||                  :::   |||      ::::::
Db     175288 TCAGCCGCAAAAATCGCTGAAGCAACTCCAGAACAAACAAAACGATTACGAATTTTTGCA 175347

Qy        208 TyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGlyAspGlu 227
              :::::::|||:::   |||||||||   |||||||::||||||      |||||||||
Db     175348 GAAAATATCGGCATTGGTTTTCAAATTAGCGACGATATTTTAGATGTAATTGGCGATGAA 175407

Qy        228 AlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyrValSer 247
              |||:::|||||||||   |||   |||         ||||||||||||||
Db     175408 ACAAAAATGGGTAAAAAGACAGGGGTCGACGCTTTTCTGAATAAAAGTACCTATCCCGGA 175467

Qy        248 LeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAlaValAsp 267
              ||||||      :::||||||:::      |||      |||         |||
Db     175468 TTACTCACACTCGAAGGTGCTAAACGGGCATTAAATGAGCATGTTTCAATTGCAAAGTCA 175527

Qy        268 GluLeuThrGlnIleAspGluGlnPheAsnThrLysHisLeuLeuGluIleValAspLeu 287
              |||:::      |||      |||:::   :::   ||||||:::::   ||||||
Db     175528 GCGCTTTCAGGACATGAT------TTCGACGATGAAATTCTCCTTAAACTTGCTGATTTA 175581

RESULT 14
AX417044
LOCUS       AX417044               349980 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 4035 from Patent WO0228891.

ACCESSION   AX417044
VERSION     AX417044.1  GI:21449654
KEYWORDS    .
SOURCE      Listeria innocua.
  ORGANISM  Listeria innocua
            Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
REFERENCE   1
  AUTHORS   Glaser,P. and Kunst,F.
  TITLE     Listeria innocua, genome and applications
  JOURNAL   Patent: WO 0228891-A 4035 11-APR-2002;
            Pasteur Institut (FR)
FEATURES             Location/Qualifiers
     source          1. .349980
                     /organism="Listeria innocua"
                     /db_xref="taxon:1642"
                     /note="seq 2058, original length: 3.011.208 replaced
                     by-seq 2058: 0.000.001 to 0.349.980-seq 4032: 0.300.001 to
                     0.649.980-seq 4033: 0.600.001 to 0.949.980-seq 4034:
                     0.900.001 to 1.249.980-seq 4035: 1.200.001 to
                     1.549.980-seq 4036: 1.500.001 to 1.849.980-seq 4037:
                     1.800.001 to 2.149.980-seq 4038: 2.100.001 to
                     2.449.980-seq 4039: 2.400.001 to 2.749.980-seq 4040:
                     2.700.001 to 3.049.980-seq 4041: 3.000.001 to 3.011.208"
BASE COUNT   116021 a  60232 c  66706 g 107021 t
ORIGIN

Alignment Scores:
Pred. No.:                2.77e-47       Length:          349980
Score:                    686.50         Matches:         146
Percent Similarity:       68.21%         Conservative:    45
Best Local Similarity:    52.14%         Mismatches:      84
Query Match:              47.25%         Indels:          5
DB:                       6              Gaps:            2

US-09-925-637-64 (1-287) x AX417044 (1-349980)

Qy          8 LysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerValMetAsp 27
              |||:::::::||||||            |||         ||||||         ::::::
Db     192757 AAAGTACTTGATGAG---------TCGCTTTTTAAAGAAATAAATATGCGCAATATCGAA 192807

Qy         28 ThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgPro 47
              :::|||:::||||||||||||||||||:::   ||||||||||||||||||||||||||
Db     192808 CCTAAGTTGAAAGAATCAATGTTATATTCAGTGCAAGCTGGCGGAAAACGAATTCGTCCA 192867

Qy         48 ValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSer 67
              :::|||:::      ||||||   :::||||||  |||         |||:::|||:::
Db     192868 ATGCTTGTTTTTGCAACGCTTCAAGCCTTAAATATTGAGCCAATGCGCGGTTTAAAAACA 192927

Qy         68 AlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMet 87
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     192928 GCTACGGCACTTGAAATGATTCATACGTACAGCTTAATTCACGATGATTTACCAGCAATG 192987

Qy         88 AspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrpThr 107
              |||:::||||||||||||||||||   ||||||||||:::|||||::   |||
Db     192988 GATGATGATGACTATCGACGCGGGAAATGGACGAACCATAAAATATACGGTGATGCAACA 193047

Qy        108 AlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAsp 127
              |||||||||||||||||||||||||||||   ||||||   :::::::::   |||:::
Db     193048 GCAATTTTAGCAGGAGATGCTTTGTTAACACTCGCTTTTTCTATTTTAGCTGAAGATGAA 193107

Qy        128 ArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHis 147
              |||:::   |||   :::|||   ::::::   :::::::|||   :::||||||
Db     193108 AATCTATCTTTCGAAACGCGTATTGCTTTAATTAATCAAATTAGTTATAGCAGTGGAGCA 193167

Qy        148 ValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIleAspLeu 167
              ||||||||||||||||||||      ||||||::::::|||   :::   :::   |||
Db     193168 GAAGGCATGGTAGGTGGGCAACAAGCCGATATGGAAGCTGAAAATAAACAAGTCACATTA 193227

Qy        168 GluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMet 187
              |||   |||      ||||||      |||||||||   ||||||   |||||||||
Db     193228 GAAGAACTAGCATCAATCCACGCTCGCAAAACTGGTGAACTATTAATTTTCGCAGTAACC 193287

Qy        188 SerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyrSer 207
```

```
Db     193288 TCAGCCGCAAAAATCGCTGAAGCAACTCCAGAACAAACAAAACGATTACGAATTTTTGCA 193347

Qy        208 TyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGlyAspGlu 227

Db     193348 GAAAATATCGGCATTGGTTTTCAAATTAGCGACGATATTTTAGATGTAATTGGCGATGAA 193407

Qy        228 AlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyrValSer 247

Db     193408 ACAAAAATGGGTAAAAAGACAGGGGTCGACGCTTTTCTGAATAAAAGTACCTATCCCGGA 193467

Qy        248 LeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAlaValAsp 267

Db     193468 TTACTCACACTCGAAGGTGCTAAACGGGCATTAAATGAGCATGTTTCAATTGCAAAGTCA 193527

Qy        268 GluLeuThrGlnIleAspGluGlnPheAsnThrLysHisLeuLeuGluIleValAspLeu 287

Db     193528 GCGCTTTCAGGACATGAT------TTCGACGATGAAATTCTCCTTAAACTTGCTGATTTA 193581

RESULT 15
E11868
LOCUS       E11868                   894 bp    DNA     linear   PAT 29-SEP-1997
DEFINITION  Mutated DNA encoding farnesyl-2-phosphate synthase.
ACCESSION   E11868
VERSION     E11868.1  GI:22025489
KEYWORDS    JP 1996214877-A/4.
SOURCE      Bacillus stearothermophilus.
  ORGANISM  Geobacillus stearothermophilus
            Bacteria; Firmicutes; Bacillales; Geobacillus.
REFERENCE   1  (bases 1 to 894)
  AUTHORS   Koike,A., Obata,A., Nishino,T., Onuma,S., Nakazawa,T., Ogura,K. and
            Furuyama,T.
  TITLE     MUTANT FARNESYL DIPHOSPHATE SYNTHETASE CAPABLE OF SYNTHESIZING
            GERANYLGERANYL DIPHOSPHATE AND DNA CODING THE ENZYME
  JOURNAL   Patent: JP 1996214877-A 4 27-AUG-1996;
            TOYOTA MOTOR CORP
COMMENT     OS   Bacillus stearothermophilus
            PN   JP 1996214877-A/4
            PD   27-AUG-1996
            PF   14-FEB-1995 JP 1995025253
            PI   KOIKE AYUMI, OBATA ATSUO, NISHINO TOKUZO, ONUMA SHINICHI, PI
            NAKAZAWA TAKESHI, OGURA KYOZO, FURUYAMA TANETOSHI PC
            C12N9/12,C07H21/04,C12N1/21,C12N15/09,C12P11/00//G01N33/50, PC
            (C12N9/12,
            PC   C12R1:19),(C12N1/21,C12R1:19),(C12N15/09,C12R1:07); CC
            strandedness: Double;
            CC   topology: Linear;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .894
            FT                   /organism='Bacillus stearothermophilus'.
FEATURES             Location/Qualifiers
     source          1. .894
                     /organism="Geobacillus stearothermophilus"
                     /db_xref="taxon:1422"
BASE COUNT      207 a    243 c    274 g    170 t
ORIGIN

Alignment Scores:
Pred. No.:              5.95e-48      Length:         894
Score:                  657.00        Matches:        133
Percent Similarity:     68.10%        Conservative:   57
Best Local Similarity:  47.67%        Mismatches:     83
Query Match:            45.22%        Indels:         6
DB:                     6             Gaps:           2

US-09-925-637-64 (1-287) x E11868 (1-894)

Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20

Db          1 ATGGCGCAGCTTTCAGTTGAACAGTTTCTCAACGAGCAAAAACAGGCGGTGGAAACAGCG 60

Qy         21 IleAsnLysSerValMet----------AspThrGlnLeuGluGluSerMetLeuTyr 36

Db         61 CTCTCCCGTTATATAGAGCGCTTAGAAGGGCCGGCGAAGCTGAAAAAGGCGATGGCGTAC 120

Qy         37 SerLeuAsnAlaGlyGlyLysArgIleArgProValLeuLeuLeuThrLeuAspSer 56

Db        121 TCATTGGAGGCCGGCGGCAAACGAATCCGTCCGTTGCTGCTTCTGTCCACCGTTCGGGCG 180

Qy         57 LeuAsnThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThr 76

Db        181 CTCGGCAAAGACCCGGCGGTCGGATTGCCCGTCGCCTGCGCGATTGAAATGATCCATACG 240

Qy         77 TyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLys 96

Db        241 CACTCTTTGATCCATGATGATTTGCCGAGCATGGACAACGATGATTTGCGGCGCGGCAAG 300

Qy         97 LeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeu 116

Db        301 CCGACGAACCATAAAGTGTTCGGCGAGGCGATGGCCATCTTGGCGGGGGACGGGTTGTTG 360

Qy        117 ThrLysAlaPheGluLeuIleSer------SerAspAspArgLeuThrAspGluValLys 134

Db        361 ACGTACGCGTTTCAATTGATCACCGAAATCGACGATGAGCGCATCCCTCCTTCCGTCCGG 420

Qy        135 IleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGln 154

Db        421 CTTCGGCTCATCGAACGGCTGGCGAAAGCGGCCGGTCCGGAAGGGATGGTCGCCGGTCAG 480

Qy        155 MetLeuAspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHis 174

Db        481 GCAGCCGATATGGAAGGAGAGGGGAAAACGCTGACGCTTTCGGAGCTCGAATACATTCAT 540

Qy        175 LysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsn 194

Db        541 CGGCATAAAACCGGGAAAATGCTGCAATACAGCGTGCACGCCGGCGCCTTGATCGGCGGC 600

Qy        195 ValAspAspThrThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPhe 214

Db        601 GCTGATGCCCGGCAAACGCGGGAGCTTGACGAATTCGCCGCCCATCTAGGCCTTGCCTTT 660

Qy        215 GlnIleLysAspAspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysVal 234

Db        661 CAAATTCGCGATGATATTCTCGATATTGAAGGGGCAGAAGAAAAAATCGGCAAGCGGGTC 720

Qy        235 GlySerAspLeuGluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAla 254

Db        721 GGCAGCGACCAAAGCAACAACAAAGCGACGTATCCAGCGTTGCTGTCGCTTGCCGGCGCG 780

Qy        255 GluAspLysLeuThrTyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAsp 273

Db        781 AAGGAAAAGTTGACGTTCCATATCGAGGCGGCGCAGCGCCATTTACGGAACGCCGAC 837

Search completed: May 30, 2003, 21:30:48
Job time : 3210 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 30, 2003, 19:09:11 ; Search time 276 Seconds
(without alignments)
2341.752 Million cell updates/sec

Title: US-09-925-637-64
Perfect score: 1453
Sequence: 1 MTNLPMNKLIDEVNNELSVA.....ELTQIDEQFNTKHLLEIVDL 287

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09925637/runat_23052003_174817_23416/app_query.fasta_1.455
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09925637_@CGN_1_1_263_@runat_23052003_174817_23416 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		1453	100.0	861	22	AAS00821	S. aureus HGS072 e
2		1442	99.2	882	23	AAS54848	Staphylococcus aur
3		1436	98.8	882	21	AAA92031	Staphylococcus aur
4		1415	97.4	864	23	AAS51602	Staphylococcus aur
5	c	1362	93.7	1893	18	AAV74466	Staphylococcus aur
6		1015	69.9	909	24	ABN90871	Staphylococcus epi
7		1015	69.9	3444	22	AAH54780	S. epidermidis gen
8	c	1015	69.9	4045	22	AAH54565	S. epidermidis gen
9		943	64.9	783	22	AAH52415	S. epidermidis ope
10		702.5	48.3	2944528	24	ABA03041	Listeria monocytog
11		695.5	47.9	5635	24	ABQ71010	Listeria monocytog
12	c	690	47.5	413	23	AAS50245	Staphylococcus aur
13	c	686.5	47.2	495269	24	ABQ67195	Listeria innocua c
14		686.5	47.2	3011208	24	ABQ69245	Listeria innocua D
15		657	45.2	894	17	AAT40228	Mutant farnesyldip
16		656	45.1	894	17	AAT40229	Native farnesyldip
17		654	45.0	894	21	AAC63809	Bacillus stearothe
18		653	44.9	894	17	AAT40226	Mutant farnesyldip
19		650	44.7	894	17	AAT40227	Mutant farnesyldip
20		650	44.7	894	19	AAV38455	DNA encoding farne
21		650	44.7	894	19	AAV18516	Bacillus stearothe
22		650	44.7	894	20	AAX86781	Farnesyl diphospha
23		650	44.7	894	24	ABK96797	B. stearothermophi
24		647	44.5	894	14	AAQ39243	FPS DNA. Bacillus
25		645	44.4	894	17	AAT40225	Mutant farnesyldip
26		644	44.3	894	20	AAX86780	Geranyl diphosphat
27		642	44.2	894	24	ABK96808	B. stearothermophi
28		635	43.7	811	24	ABK73724	Bacillus lichenifo
29		596	41.0	882	23	AAS52896	Enterococcus faeca
30		587.5	40.4	7528	20	AAX12992	Enterococcus faeca
31	c	570	39.2	337	23	AAS50698	Staphylococcus aur
32		551.5	38.0	870	24	ABN66185	Streptococcus poly
33		551	37.9	876	21	AAA94607	Farnesyl diphospha
34		549	37.8	876	23	AAS55862	Streptococcus pneu
35	c	549	37.8	10011	19	AAV52209	Streptococcus pneu
36		547	37.6	876	23	AAS55538	Streptococcus pneu
37		540.5	37.2	2365589	24	ABA90521	Genomic sequence o
38		520.5	35.8	885	21	AAA13985	Taxus cuspidata ge
39		520.5	35.8	1179	21	AAA13994	Taxus GGPP synthas
40		520.5	35.8	1179	21	AAA13995	Taxus GGPP synthas
41		520.5	35.8	1179	21	AAA13996	Taxus GGPP synthas
42		520.5	35.8	1179	21	AAA13997	Taxus GGPP synthas
43		520.5	35.8	1179	21	AAA13998	Taxus GGPP synthas
44		520.5	35.8	1179	21	AAA13999	Taxus GGPP synthas
45		520.5	35.8	1179	21	AAA14000	Taxus GGPP synthas

ALIGNMENTS

RESULT 1
AAS00821
ID AAS00821 standard; DNA; 861 BP.
XX
AC AAS00821;
XX
DT 04-JUL-2001 (first entry)
XX
DE S. aureus HGS072 encoding Farnesyl diphosphatesynthase, IspA.
XX
KW Farnesyl diphosphatesynthase; IspA; immunogen; vaccine; antibody;
KW wound infection; cellulitis; burn infection; eyelid infection;
KW food poisoning; joint infection; neonatal conjunctivitis; osteomyelitis;
KW skin infection; scalded skin syndrome; toxic epidermal necrosis;
KW Ritter's disease; Lyell's disease; toxic shock syndrome; endocarditis;
KW HGS072; ds.
XX
OS Staphylococcus aureus.

XX
FH Key Location/Qualifiers
FT CDS 1..861
FT /*tag= a
FT /product= "IspA"
FT /partial
FT /note= "No stop codon"
XX
PN WO200116292-A2.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-US23773.
XX
PR 01-SEP-1999; 99US-0151933.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Choi GH;
XX
DR WPI; 2001-183259/18.
DR P-PSDB; AAU00859.
XX
PT New isolated nucleic acid for use in diagnosing Staphylococcus
PT infections and in vaccines for eliciting immune responses to the
PT infections -
XX
PS Claim 1; Page 23; 225pp; English.
XX
CC The sequence encodes S. aureus IspA (Farnesyl diphosphatesynthase).
CC The polynucleotides of the invention are used to detect Staphylococcus
CC nucleic acids in a biological sample from an animal for diagnosing
CC Staphylococcus infections. The polypeptides of the invention are used to
CC detect anti-Staphylococcus antibodies in a biological sample from an
CC animal to diagnose Staphylococcus infections. The polypeptides are also
CC used in vaccines to elicit protective antibodies in an animal to a member
CC of the Staphylococcus genus and for preventing or attenuating an
CC infection caused by a member of the Staphylococcus genus e.g wound
CC infection, cellulitis, burn infection, eyelid infection, food poisoning,
CC joint infection, neonatal conjunctivitis, osteomyelitis, skin infection,
CC scalded skin syndrome (also known as toxic epidermal necrosis, Ritter's
CC disease and Lyell's disease), toxic shock syndrome and endocarditis. The
CC polynucleotides may also be used in vaccines and for preventing or
CC attenuating a Staphylococcus infection. Antibodies to the polypeptides
CC may be used to purify, detect and target the polypeptides in in vitro and
CC in vivo diagnostic and therapeutic methods.
XX
SQ Sequence 861 BP; 323 A; 114 C; 180 G; 244 T; 0 other;

Alignment Scores:

Pred. No.:	1.75e-147	Length:	861
Score:	1453.00	Matches:	287
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	22	Gaps:	0

US-09-925-637-64 (1-287) x AAS00821 (1-861)

```
Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy         61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy         81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300

Qy        101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600

Qy        201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660

Qy        221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Qy        241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy        261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy        281 LeuLeuGluIleValAspLeu 287
              |||||||||||||||||||||
Db        841 TTATTAGAAATCGTTGATTTA 861
```

RESULT 2
AAS54848
ID AAS54848 standard; DNA; 882 BP.
XX
AC AAS54848;
XX
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus DNA for cellular proliferation protein #1160.
XX
KW Antisense; ds; prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.

```
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   P-PSDB; AAU36989.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Claim 27; Seq ID No 8485; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence encodes an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 882 BP; 330 A; 116 C; 185 G; 251 T; 0 other;

Alignment Scores:
Pred. No.:              2.78e-146     Length:          882
Score:                  1442.00       Matches:         285
Percent Similarity:     99.30%        Conservative:    0
Best Local Similarity:  99.30%        Mismatches:      2
Query Match:            99.24%        Indels:          0
DB:                     23            Gaps:            0

US-09-925-637-64 (1-287) x AAS54848 (1-882)

Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
            1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy         61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy         81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCAT 300

Qy        101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTATTAACATTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAA 600

Qy        201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTTCAGATTAAAGATGATTTA 660

Qy        221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTAGACTGCTATGGTGATGAAGCGAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Qy        241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATAAAAGTACATACGTGAGTTTATTAGGAAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy        261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATAGAGACGCAGCAGTGGATGAACTAATGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy        281 LeuLeuGluIleValAspLeu 287
              |||||||||||||||||||||
Db        841 TTATTAGAAATCGTTGATTTA 861

RESULT 3
AAA92031
ID   AAA92031 standard; DNA; 882 BP.
XX
AC   AAA92031;
XX
DT   12-JAN-2001  (first entry)
XX
DE   Staphylococcus aureus ispA coding sequence.
XX
KW   IspA; bacterial disease; respiratory tract infection;
KW   gastrointestinal infection; cardiac infection; Helicobacter pylori;
KW   stomach cancer; stomach ulcer; gastritis; ds.
XX
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..882
FT                   /*tag=  a
FT                   /product= "IspA"
XX
PN   US6107058-A.
XX
PD   22-AUG-2000.
XX
PF   26-MAR-1999;   99US-0276873.
XX
PR   26-MAR-1999;   99US-0276873.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI   Gwynn M,  Wilding EI;
XX
DR   WPI; 2000-578535/54.
DR   P-PSDB; AAB23333.
XX
PT   Novel farnesyl diphosphate synthase polynucleotide from staphylococcus
PT   aureus useful for diagnosis and treatment of bacterial infections and
PT   as hybridization probe for isolating genomic clones  -
```

```
XX
PS   Claim 1; column 1-4; 15pp; English.
XX
CC   The present sequence is the coding sequence for the Staphylococcus aureus
CC   IspA protein. This gene and the protein it encodes can be used in many
CC   research assays, as well as treatments for bacterial diseases such as
CC   infections of the respiratory tract (including otitis media, bacterial
CC   tracheitis, acute epiglottitis, thyroiditis, empyema and lung abscesses),
CC   cardiac infections such as infective endocarditis, gastrointestinal
CC   infections including secretory diarrhoea, splenic abscesses and
CC   retroperitoneal abscesses, CNS infections such as cerebral abscesses, eye
CC   infections (including blepharitis, conjunctivitis, keratitis,
CC   endophthalmitis, preseptal and orbital cellulitis and darcryocystitis),
CC   kidney and urinary tract infections such as epididymitis, intrarenal and
CC   perinephric abscesses and toxic shock syndrome, skin diseases (including
CC   impetigo, folliculitis, cutaneous abscesses, wound infection and
CC   bacterial myositis), bone and joint infections such as septic
CC   arthritis and osteomyelitis, septic thrombophlebitis, food poisoning and
CC   scalded skin syndrome. In addition, they can be used to treat
CC   diseases caused by Helicobacter pylori, including stomach cancer, stomach
CC   ulcers and gastritis.
XX
SQ   Sequence 882 BP; 328 A; 118 C; 186 G; 250 T; 0 other;

Alignment Scores:
Pred. No.:                1.23e-145      Length:          882
Score:                    1436.00        Matches:         284
Percent Similarity:       99.30%         Conservative:    1
Best Local Similarity:    98.95%         Mismatches:      2
Query Match:              98.83%         Indels:          0
DB:                       21             Gaps:            0

US-09-925-637-64 (1-287) x AAA92031 (1-882)

Qy      1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy     21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGCATGTTGTATTCATTAAATGCT 120

Qy     41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy     61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
          ||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db    181 TATGAGTTAGGTGTGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy     81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAAGGAAATTAACAAATCAT 300

Qy    101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy    121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy    141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy    161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy    181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db    541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAA 600

Qy    201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660

Qy    221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Qy    241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy    261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy    281 LeuLeuGluIleValAspLeu 287
          |||||||||||||||||||||
Db    841 TTATTAGAAATCGTTGATTTA 861

RESULT 4
AAS51602
ID   AAS51602 standard; DNA; 864 BP.
XX
AC   AAS51602;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus DNA for cellular proliferation protein #19.
XX
KW   Antisense; ds; prokaryotic cellular proliferation gene;
KW   antibiotic; antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   P-PSDB; AAU33743.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Claim 27; Seq ID No 4184; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
```

```
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence encodes an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 864 BP; 323 A; 112 C; 182 G; 247 T; 0 other;

Alignment Scores:
Pred. No.:              2.22e-143     Length:         864
Score:                  1415.00       Matches:        280
Percent Similarity:     99.29%        Conservative:   0
Best Local Similarity:  99.29%        Mismatches:     2
Query Match:            97.38%        Indels:         0
DB:                     23            Gaps:           0

-09-925-637-64 (1-287) x AAS51602 (1-864)

Qy          6 MetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerVal 25
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTA 60

Qy         26 MetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIle 45
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATC 120

Qy         46 ArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMet 65
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATG 180

Qy         66 LysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCA 240

Qy         86 AlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGlu 105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCATAAAGTATATGGTGAG 300

Qy        106 TrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSer 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGT 360

Qy        126 AspAspArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSer 145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          361 GATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGT 420

          146 GlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIle 165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATT 480

Qy        166 AspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAla 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCACTATTAACATTTGCG 540

Qy        186 ValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSer 205
              ||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db        541 GTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAAGAACATTTAGAAAGT 600

Qy        206 TyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGly 225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATAGTTATCATTTAGGTATGATGTTTCAGATTAAAGATGATTTATTAGACTGCTATGGT 660

Qy        226 AspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyr 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GATGAAGCGAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTACATAC 720

Qy        246 ValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAla 265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTGAGTTTATTAGGAAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGACGCAGCA 780

Qy        266 ValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHisLeuLeuGluIleVal 285
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        781 GTGGATGAACTAATGCAAATTGATGAACAATTCAATACAAAACACTTATTAGAAATCGTT 840

Qy        286 AspLeu 287
              ||||||
Db        841 GATTTA 846

RESULT 5
AAV74466/c
ID   AAV74466 standard; DNA; 1893 BP.
XX
AC   AAV74466;
XX
DT   16-MAR-1999  (first entry)
XX
DE   Staphylococcus aureus contig SEQ ID #155.
XX
KW   Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW   cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW   skin infection; surgical wound infection; scalded skin syndrome;
KW   toxic shock syndrome; ds.
XX
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1261..1320
FT                   /*tag= a
FT                   /note= "these bases represent a line of missing text in
FT                          the sequence listing in the specification. They
FT                          are included to maintain the nucleotide numbering
FT                          given in the specification for this DNA sequence"
XX
PN   EP786519-A2.
XX
PD   30-JUL-1997.
XX
PF   07-JAN-1997;   97EP-0100117.
XX
PR   05-JAN-1996;   96US-0009861.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Barash SC,  Choi GH,  Dillon PJ,  Fannon MR,  Kunsch CA;
PI   Rosen CA;
XX
DR   WPI; 1997-374922/35.
XX
PT   Polynucleotide(s) and proteins derived from Staphylococcus aureus -
PT   stored on computer readable medium and used in the production of
PT   anti-S.aureus vaccines
XX
PS   Claim 1; Page 784-785; 3271pp; English.
XX
CC   This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC   of the invention. The DNA sequences are recorded on a computer readable
CC   medium, preferably selected from a floppy or hard disk, random access
CC   memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC   the S.aureus DNA sequences allows putative functions to be assigned so
CC   that protein-encoding or regulatory regions of commercial, therapeutic or
CC   industrial importance can be obtained. Specifically, sequences which are
CC   likely to encode antigens have been identified and these polypeptides can
CC   be used in a vaccine composition against S.aureus infection. The
CC   polypeptides can also be used in a kit for the immunodetection of
CC   S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC   including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC   skin and surgical wound infections, scalded skin syndrome, toxic shock
CC   syndrome, etc. Organisms transformed with the DNA sequences can be used
CC   for recombinant production of the polypeptides. The new DNA sequences
CC   (and their fragments) are useful as primers or probes for isolating
CC   homologues of any of the S.aureus DNA sequences contained on the
CC   computer readable medium.
```

```
XX
SQ   Sequence 1893 BP; 552 A; 336 C; 233 G; 708 T; 64 other;

Alignment Scores:
Pred. No.:                3.26e-137      Length:          1893
Score:                    1362.00        Matches:         269
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              93.74%         Indels:          0
DB:                       18             Gaps:            0

US-09-925-637-64 (1-287) x AAV74466 (1-1893)

Qy         19 ValAlaIleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeu 38
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GTTGCGATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTA 1201

Qy         39 AsnAlaGlyGlyLysArgIleArgProValLeuLeuLeuThrLeuAspSerLeuAsn 58
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 AATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAAT 1141

Qy         59 ThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSer 78
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 ACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCA 1081

Qy         79 LeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThr 98
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 CTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACA 1021

Qy         99 AsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLys 118
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 AATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAA 961

Qy        119 AlaPheGluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeu 138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 GCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTA 901

Qy        139 GlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMet 158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 CAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATG 841

Qy        159 GlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThr 178
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 CAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACA 781

Qy        179 GlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThr 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
            780 GGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACA 721

            199 ThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAsp 218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ACTAAAGAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGAT 661

Qy        219 AspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeu 238
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 GATTTATTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTT 601

Qy        239 GluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeu 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GAAAATAATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTG 541

Qy        259 ThrTyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThr 278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 ACTTATCATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACA 481

Qy        279 LysHisLeuLeuGluIleValAspLeu 287
              |||||||||||||||||||||||||||
Db        480 AAACACTTATTAGAAATCGTTGATTTA 454

RESULT 6
ABN90871
ID   ABN90871 standard; DNA; 909 BP.
XX
AC   ABN90871;
XX
DT   24-JUL-2002  (first entry)
XX
DE   Staphylococcus epidermidis ORF nucleic acid sequence SEQ ID NO:334.
XX
KW   Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW   antibacterial; gene therapy; gene; ds.
XX
OS   Staphylococcus epidermidis.
XX
PN   US6380370-B1.
XX
PD   30-APR-2002.
XX
PF   13-AUG-1998;   98US-0134001.
XX
PR   14-AUG-1997;   97US-055779P.
PR   08-NOV-1997;   97US-064964P.
XX
PA   (GENO-) GENOME THERAPEUTICS CORP.
XX
PI   Doucette-Stamm LA,  Bush D;
XX
DR   WPI; 2002-381255/41.
DR   P-PSDB; ABP38326.
XX
PT   Novel isolated nucleic acid encoding a Staphylococcus epidermidis
PT   polypeptide, useful for diagnosing and treating bacterial infections -
XX
PS   Disclosure; SEQ ID 334; 267pp; English.
XX
CC   ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC   frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC   given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC   antibacterial activity and can be used in gene therapy. The sequences
CC   can also be used in the diagnosis and treatment of bacterial infections,
CC   particularly S. epidermidis infections. The sequences can be used to
CC   screen for compounds able to interfere with the S. epidermidis life
CC   cycle or inhibit S. epidermidis infection.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the
CC   USPTO web site.
XX
SQ   Sequence 909 BP; 361 A; 106 C; 165 G; 277 T; 0 other;

Alignment Scores:
Pred. No.:                3.61e-100      Length:          909
Score:                    1015.00        Matches:         197
Percent Similarity:       82.93%         Conservative:    41
Best Local Similarity:    68.64%         Mismatches:      49
Query Match:              69.86%         Indels:          0
DB:                       24             Gaps:            0

US-09-925-637-64 (1-287) x ABN90871 (1-909)

Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              |||       |||   |||||||||||||||:::   :::|||      |||:::   :::
Db         28 ATGAAGAAACTACAGATGAATAAATTAATAAATATAATTAATACATCACTGAATAAGTCA 87

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              |||       |||   :::   |||   ||||||||||||||   ||||||||||||||
Db         88 ATACAATCATCACCATTAAAAACTAATTTAGAAGAAAGTATGAAATATTCATTAAATGCT 147

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||||:::|||||||||||||||      ||||||   :::
Db        148 GGTGGTAAAAGAATCAGACCAGTCATATTATTATTAACACTAAAAATGCTTAACAAAGAT 207

Qy         61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
              |||:::   |||:::   ||||||:::||||||||||||||||||||||||||||||||
Db        208 TATCAACAAGGACTAAATAGTGCTTTAGCATTGGAAATGATTCATACTTATTCTTTAATT 267

Qy         81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CATGATGATTTACCAGCAATGGATAATGACGATTACCGTAGAGGAAAATTAACAAATCAT 327

Qy        101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
                |||||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db        328 AAAGTTTATGGTGAATGGAAAGCCATTCTTGCTGGTGATGCATTATTAACAAAAGCTTTT 387

Qy        121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
                ||||||:::|||:::|||      :::   |||   ||||||:::   :::::::::|||
Db        388 GAATTAGTTTCTAATGATACTACCATTGAAGATAGTGTGAAAGTAAGTATTATAAAAAGA 447

Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
                ||||||   ||||||||||||:::||||||||||||||||||   |||||||||:::|||
Db        448 CTTTCAAAAGCAAGTGGACATTTGGGAATGGTGGGTGGCCAAGCGCTTGATATGGAAAGT 507

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
                ||||||:::   |||   ||||||||||||||   ||||||:::||||||||||||||||
Db        508 GAAGGGAAGTCAATTCGTTTAGAAACTTTAGAATCAATTCATGAAACTAAGACAGGCGCT 567

Qy        181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys 200
                ||||||   |||:::||||||:::|||   |||||||||   |||:::
Db        568 TTACTAAATTTTTCAGTTATGGCTGCGGTAGACATTGCTCAAGTAGAACAAAATATTGCT 627

Qy        201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
                ::::::|||:::   :::|||:::|||||||||||||||||||||||||||||||||||||
Db        628 AAGAATTTAGATGAATTTAGTCATCATTTAGGAATGATGTTTCAAATTAAAGATGATTTA 687

Qy        221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
                ||||||   ||||||||||||:::||||||||||||||||||||||||||:::   |||
Db        688 CTGGATGTGTATGGTGATGAATCAAAACTTGGCAAAAAAGTAGGCAGTGATATAGTAAAT 747

Qy        241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
                :::|||||||||||||||||||||||||||||:::|||||||||:::||||||
Db        748 CATAAAAGTACTTATGTTTCTTTACTTGGAAAAGAAGGAGCAGAAGAAAAGTTAAACAAT 807

Qy        261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
                |||:::      |||::::::   |||   ||||||   :::|||::::::|||
Db        808 CATCAATATCTTGCTATGAACTGCTTAAATCAAATTTCTGATCAATATGATACTTCTGAA 867

Qy        281 LeuLeuGluIleValAspLeu 287
                |||   :::|||||||||||||
Db        868 TTAAGTGATATTGTAGATTTA 888

RESULT 7
AAH54780
ID   AAH54780 standard; DNA; 3444 BP.
XX
AC   AAH54780;
XX
DT   03-SEP-2001  (first entry)
XX
DE   S. epidermidis genomic polynucleotide sequence SEQ ID NO:4144.
XX
KW   Staphylococcus epidermidis SR1 strain; infection; diagnosis;
KW   vaccination; endocarditis; ds.
XX
OS   Staphylococcus epidermidis.
XX
PN   WO200134809-A2.
XX
PD   17-MAY-2001.
XX
PF   09-NOV-2000; 2000WO-US30782.
XX
PR   09-NOV-1999;   99US-0164258.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Kimmerly WJ;
XX
DR   WPI; 2001-316495/33.
XX
PT   Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT   useful for vaccinating against infections, e.g. endocarditis -
XX
PS   Claim 8; Page 1844-1845; 2188pp; English.
XX
CC   AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC   (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC   (I) and (II) can have antibacterial activity and therefore can be used
CC   in vaccination. The nucleic acids (I) may be used to produce the
CC   S. epidermidis polypeptides (II) via the production of vectors
CC   containing them which are used to produce hosts cells which express the
CC   polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC   used to vaccinate subjects and to raise antibodies against the bacteria.
CC   The polypeptides may also be used to assay for other inhibitors of their
CC   activity and therefore identify compounds that may be used for the
CC   treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC   AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC   polynucleotide sequences from the present invention. AAH55091 to
CC   AAH55098 represent oligonucleotide sequences and primers which are used
CC   in the exemplification of the present invention.
CC   N.B. The present invention specifically claims all the polynucleotide
CC   sequences given in the sequence listing of the present specification,
CC   however the sequence listing only goes up to SEQ ID NO:4454 so even
CC   though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC   no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ   Sequence 3444 BP; 1365 A; 451 C; 571 G; 1057 T; 0 other;

Alignment Scores:
Pred. No.:                2.06e-99      Length:          3444
Score:                    1015.00       Matches:         197
Percent Similarity:       82.93%        Conservative:    41
Best Local Similarity:    68.64%        Mismatches:      49
Query Match:              69.86%        Indels:          0
DB:                       22            Gaps:            0

US-09-925-637-64 (1-287) x AAH54780 (1-3444)

Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
                |||       |||   |||||||||||||||:::   :::|||      |||:::   :::
Db        766 ATGAAGAAACTACAGATGAATAAATTAATAAATATAATTAATACATCACTGAATAAGTCA 825

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
                |||       |||   :::   |||   ||||||||||||||   ||||||||||||||
Db        826 ATACAATCATCACCATTAAAAACTAATTTAGAAGAAAGTATGAAATATTCATTAAATGCT 885

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
                ||||||||||||||||||||||||:::|||||||||||||||      ||||||   :::
Db        886 GGTGGTAAAAGAATCAGACCAGTCATATTATTATTAACACTAAAAATGCTTAACAAAGAT 945

Qy         61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
                |||:::   |||:::   ||||||:::||||||||||||||||||||||||||||||||||
Db        946 TATCAACAAGGACTAAATAGTGCTTTAGCATTGGAAATGATTCATACTTATTCTTTAATT 1005

Qy         81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1006 CATGATGATTTACCAGCAATGGATAATGACGATTACCGTAGAGGAAAATTAACAAATCAT 1065

Qy        101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
                |||||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db       1066 AAAGTTTATGGTGAATGGAAAGCCATTCTTGCTGGTGATGCATTATTAACAAAAGCTTTT 1125

Qy        121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
                ||||||:::|||:::|||      :::   |||   |||||||:::   :::::::::|||
Db       1126 GAATTAGTTTCTAATGATACTACCATTGAAGATAGTGTGAAAGTAAGTATTATAAAAAGA 1185

Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
                ||||||   ||||||||||||:::||||||||||||||||||   |||||||||:::|||
Db       1186 CTTTCAAAAGCAAGTGGACATTTGGGAATGGTGGGTGGCCAAGCGCTTGATATGGAAAGT 1245

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
                ||||||:::   |||   ||||||||||||||   ||||||:::||||||||||||||||
Db       1246 GAAGGGAAGTCAATTCGTTTAGAAACTTTAGAATCAATTCATGAAACTAAGACAGGCGCT 1305
```

```
Qy        181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys 200
              ||||||   |||:::||||||:::|||   |||||||||   |||:::
Db       1306 TTACTAAATTTTTCAGTTATGGCTGCGGTAGACATTGCTCAAGTAGAACAAAATATTGCT 1365

Qy        201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
              ::::::|||:::   :::|||:::|||||||||||||||||||||||||||||||||||||
Db       1366 AAGAATTTAGATGAATTTAGTCATCATTTAGGAATGATGTTTCAAATTAAAGATGATTTA 1425

Qy        221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
              ||||||   ||||||||||||:::|||||||||||||||||||||||||:::   |||
Db       1426 CTGGATGTGTATGGTGATGAATCAAAACTTGGCAAAAAAGTAGGCAGTGATATAGTAAAT 1485

Qy        241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
              :::||||||||||||||||||||||||||||||:::||||||||||:::|||||||
Db       1486 CATAAAAGTACTTATGTTTCTTTACTTGGAAAAGAAGGAGCAGAAGAAAAGTTAAACAAT 1545

Qy        261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
              |||:::      |||:::::::   |||   ||||||   :::|||::::::|||
Db       1546 CATCAATATCTTGCTATGAACTGCTTAAATCAAATTTCTGATCAATATGATACTTCTGAA 1605

Qy        281 LeuLeuGluIleValAspLeu 287
              |||   :::||||||||||||
Db       1606 TTAAGTGATATTGTAGATTTA 1626
```

RESULT 8
AAH54565/c
ID AAH54565 standard; DNA; 4045 BP.
XX
AC AAH54565;
XX
DT 03-SEP-2001 (first entry)
XX
DE S. epidermidis genomic polynucleotide sequence SEQ ID NO:3929.
XX
KW Staphylococcus epidermidis SR1 strain; infection;. diagnosis;
KW vaccination; endocarditis; ds.
XX
OS Staphylococcus epidermidis.
XX
PN WO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US30782.
XX
PR 09-NOV-1999; 99US-0164258.
XX
(GLAX) GLAXO GROUP LTD.

Kimmerly WJ;
XX
DR WPI; 2001-316495/33.
XX
PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis -
XX
PS Claim 8; Page 1586-1588; 2188pp; English.
XX
CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ Sequence 4045 BP; 1249 A; 691 C; 536 G; 1569 T; 0 other;

Alignment Scores:

Pred. No.:	2.54e-99	Length:	4045
Score:	1015.00	Matches:	197
Percent Similarity:	82.93%	Conservative:	41
Best Local Similarity:	68.64%	Mismatches:	49
Query Match:	69.86%	Indels:	0
DB:	22	Gaps:	0

US-09-925-637-64 (1-287) x AAH54565 (1-4045)

```
Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              |||      |||  ||||||||||||||||:::  :::|||      |||:::  :::
Db       2313 ATGAAGAAACTACAGATGAATAAATTAATAAATATAATTAATACATCACTGAATAAGTCA 2254

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              |||      |||  :::  |||  ||||||||||||||  ||||||||||||||
Db       2253 ATACAATCATCACCATTAAAAACTAATTTAGAAGAAAGTATGAAATATTCATTAAATGCT 2194

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||:::|||||||||||||||  ||||||  :::
Db       2193 GGTGGTAAAAGAATCAGACCAGTCATATTATTATTAACACTAAAAATGCTTAACAAAGAT 2134

Qy         61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
              |||:::  |||:::  ||||||:::|||||||||||||||||||||||||||||||||||
Db       2133 TATCAACAAGGACTAAATAGTGCTTTAGCATTGGAAATGATTCATACTTATTCTTTAATT 2074

Qy         81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2073 CATGATGATTTACCAGCAATGGATAATGACGATTACCGTAGAGGAAAATTAACAAATCAT 2014

Qy        101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
              ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db       2013 AAAGTTTATGGTGAATGGAAAGCCATTCTTGCTGGTGATGCATTATTAACAAAAGCTTTT 1954

Qy        121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
              ||||||:::|||:::|||      :::  |||  |||||||:::  :::::::::|||
Db       1953 GAATTAGTTTCTAATGATACTACCATTGAAGATAGTGTGAAAGTAAGTATTATAAAAAGA 1894

Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
              ||||||   ||||||||||||:::|||||||||||||||||||  ||||||||||:::|||
Db       1893 CTTTCAAAAGCAAGTGGACATTTGGGAATGGTGGGTGGCCAAGCGCTTGATATGGAAAGT 1834

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
              ||||||:::   |||  ||||||||||||||  ||||||:::||||||||||||||||
Db       1833 GAAGGGAAGTCAATTCGTTTAGAAACTTTAGAATCAATTCATGAAACTAAGACAGGCGCT 1774

Qy        181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys 200
              ||||||   |||:::|||||||:::|||   |||||||||   |||:::
Db       1773 TTACTAAATTTTTCAGTTATGGCTGCGGTAGACATTGCTCAAGTAGAACAAAATATTGCT 1714

Qy        201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
              ::::::|||:::   :::|||:::|||||||||||||||||||||||||||||||||||||
Db       1713 AAGAATTTAGATGAATTTAGTCATCATTTAGGAATGATGTTTCAAATTAAAGATGATTTA 1654

Qy        221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
              ||||||   ||||||||||||:::|||||||||||||||||||||||||:::   |||
Db       1653 CTGGATGTGTATGGTGATGAATCAAAACTTGGCAAAAAAGTAGGCAGTGATATAGTAAAT 1594

Qy        241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
              :::||||||||||||||||||||||||||||||:::||||||||||:::|||||||
Db       1593 CATAAAAGTACTTATGTTTCTTTACTTGGAAAAGAAGGAGCAGAAGAAAAGTTAAACAAT 1534

Qy        261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
              |||:::      |||::::::   |||   ||||||   :::|||::::::|||
Db       1533 CATCAATATCTTGCTATGAACTGCTTAAATCAAATTTCTGATCAATATGATACTTCTGAA 1474
```

```
Qy        281 LeuLeuGluIleValAspLeu 287
              |||   :::|||||||||||||
Db       1473 TTAAGTGATATTGTAGATTTA 1453

RESULT 9
AAH52415
ID   AAH52415 standard; DNA; 783 BP.
XX
AC   AAH52415;
XX
DT   03-SEP-2001  (first entry)
XX
DE   S. epidermidis open reading frame nucleotide sequence SEQ ID NO:223.
XX
KW   Staphylococcus epidermidis SR1 strain; infection; diagnosis;
KW   vaccination; endocarditis; ds.
XX
OS   Staphylococcus epidermidis.
XX
PN   WO200134809-A2.
XX
PD   17-MAY-2001.
XX
PF   09-NOV-2000; 2000WO-US30782.
XX
PR   09-NOV-1999;   99US-0164258.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Kimmerly WJ;
XX
DR   WPI; 2001-316495/33.
DR   P-PSDB; AAG81565.
XX
PT   Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT   useful for vaccinating against infections, e.g. endocarditis -
XX
PS   Claim 8; Page 103; 2188pp; English.
XX
CC   AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC   (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC   (I) and (II) can have antibacterial activity and therefore can be used
CC   in vaccination. The nucleic acids (I) may be used to produce the
CC   S. epidermidis polypeptides (II) via the production of vectors
CC   containing them which are used to produce hosts cells which express the
CC   polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC   used to vaccinate subjects and to raise antibodies against the bacteria.
CC   The polypeptides may also be used to assay for other inhibitors of their
CC   activity and therefore identify compounds that may be used for the
CC   treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC   AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC   polynucleotide sequences from the present invention. AAH55091 to
CC   AAH55098 represent oligonucleotide sequences and primers which are used
CC   in the exemplification of the present invention.
CC   N.B. The present invention specifically claims all the polynucleotide
CC   sequences given in the sequence listing of the present specification,
CC   however the sequence listing only goes up to SEQ ID NO:4454 so even
CC   though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC   no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ   Sequence 783 BP; 297 A; 93 C; 147 G; 246 T; 0 other;

Alignment Scores:
Pred. No.:              1.76e-92      Length:          783
Score:                  943.00        Matches:         181
Percent Similarity:     85.43%        Conservative:    36
Best Local Similarity:  71.26%        Mismatches:      37
Query Match:            64.90%        Indels:          0
DB:                     22            Gaps:            0

US-09-925-637-64 (1-287) x AAH52415 (1-783)

Qy         34 MetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgProValLeuLeuLeuLeuThr 53
              |||   |||||||||||||||||||||||||||||||||||||:::|||||||||||||
Db          1 ATGAAATATTCATTAAATGCTGGTGGTAAAAGAATCAGACCAGTCATATTATTATTAACA 60

Qy         54 LeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMet 73
              |||      ||||||   :::|||:::   |||:::   ||||||:::|||||||||||||
Db         61 CTAAAAATGCTTAACAAAGATTATCAACAAGGACTAAATAGTGCTTTAGCATTGGAAATG 120

Qy         74 IleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArg 93
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATTCATACTTATTCTTTAATTCATGATGATTTACCAGCAATGGATAATGACGATTACCGT 180

Qy         94 ArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAsp 113
              |||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db        181 AGAGGAAAATTAACAAATCATAAAGTTTATGGTGAATGGAAAGCCATTCTTGCTGGTGAT 240

Qy        114 AlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAspArgLeuThrAspGluVal 133
              |||||||||||||||||||||||||||:::|||:::|||      :::   |||   |||
Db        241 GCATTATTAACAAAAGCTTTTGAATTAGTTTCTAATGATACTACCATTGAAGATAGTGTG 300

Qy        134 LysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGly 153
              |||:::   :::::::::|||||||||   ||||||||||||:::|||||||||||||||
Db        301 AAAGTAAGTATTATAAAAAGACTTTCAAAAGCAAGTGGACATTTGGGAATGGTGGGTGGC 360

Qy        154 GlnMetLeuAspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIle 173
              |||   |||||||||:::||||||||:::   |||   ||||||||||||||   |||
Db        361 CAAGCGCTTGATATGGAAAGTGAAGGGAAGTCAATTCGTTTAGAAACTTTAGAATCAATT 420

Qy        174 HisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAla 193
              |||:::||||||||||||||||||||   |||:::|||||:::|||   |||||||||
Db        421 CATGAAACTAAGACAGGCGCTTTACTAAATTTTTCAGTTATGGCTGCGGTAGACATTGCT 480

Qy        194 AsnValAspAspThrThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMet 213
                 |||:::           :::::::|||:::   :::|||:::||||||||||||||
Db        481 CAAGTAGAACAAAATATTGCTAAGAATTTAGATGAATTTAGTCATCATTTAGGAATGATG 540

Qy        214 PheGlnIleLysAspAspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLys 233
              ||||||||||||||||||||||||||   |||||||||||:::|||||||||||||||||
Db        541 TTTCAAATTAAAGATGATTTACTGGATGTGTATGGTGATGAATCAAAACTTGGCAAAAAA 600

Qy        234 ValGlySerAspLeuGluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGly 253
              ||||||||||||::   |||:::|||||||||||||||||||||||||||:::|||
Db        601 GTAGGCAGTGATATAGTAAATCATAAAAGTACTTATGTTTCTTTACTTGGAAAAGAAGGA 660

Qy        254 AlaGluAspLysLeuThrTyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAsp 273
              ||||||:::||||||      |||:::      |||::::::   |||   ||||||
Db        661 GCAGAAGAAAAGTTAAACAATCATCAATATCTTGCTATGAACTGCTTAAATCAAATTTCT 720

Qy        274 GluGlnPheAsnThrLysHisLeuLeuGluIleValAspLeu 287
              :::|||:::::::|||      |||   :::|||||||||||||
Db        721 GATCAATATGATACTTCTGAATTAAGTGATATTGTAGATTTA 762

RESULT 10
ABA03041
ID   ABA03041 standard; DNA; 2944528 BP.
XX
AC   ABA03041;
XX
DT   05-FEB-2002  (first entry)
XX
DE   Listeria monocytogenes EGD-e genome sequence.
XX
KW   Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW   vitamin B12; bacterial infection; disease; ds.
XX
OS   Listeria monocytogenes.
XX
PN   WO200177335-A2.
XX
PD   18-OCT-2001.
XX
PF   11-APR-2001; 2001WO-FR01118.
XX
```

```
PR   11-APR-2000; 2000FR-0004629.
XX
PA   (INSP ) INST PASTEUR.
XX
PI   Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
PI   Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI   Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
PI   Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI   Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
PI   Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI   Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
PI   Rose M, Voss H;
XX
DR   WPI; 2002-010914/01.
XX
PT   Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT   and prevention of Listeria and related bacterial infections, and
PT   related polypeptides -
XX
PS   Claim 1; SEQ ID No 1; 192pp; French.

CC   The present sequence is the genome sequence of Listeria monocytogenes
CC   EGD-e. This sequence and fragments of this sequence are useful for
CC   selecting probes and primers for detecting genes in L. monocytogenes and
CC   related organisms, and to study genetic polymorphisms and other genomes.
CC   Proteins (ABB47297-ABB50149) expressed from the present sequence are
CC   useful for raising specific antibodies, identification of L.
CC   monocytogenes and related organisms, and for biosynthesis and
CC   biodegradation, especially biosynthesis of Vitamin B12. This sequence and
CC   proteins encoded by it are also useful for selecting compounds that
CC   regulate gene expression and cell replication and modulate L.
CC   monocytogenes-related diseases. In addition, this sequence and proteins
CC   encoded by it are useful in pharmaceutical and vaccines compositions for
CC   the treatment or prevention of infections by L. monocytogenes and related
CC   organisms.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2944528 BP; 914202 A; 563301 C; 555061 G; 911964 T; 0 other;

Alignment Scores:
Pred. No.:              7.57e-62        Length:         2944528
Score:                  702.50          Matches:        148
Percent Similarity:     69.29%          Conservative:   46
Best Local Similarity:  52.86%          Mismatches:     81
Query Match:            48.35%          Indels:         5
[illegible]:            24              Gaps:           2

-09-925-637-64 (1-287) x ABA03041 (2944528)

Qy          8 LysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerValMetAsp 27
              |||:::|||||||||            |||       ||||||:::      ::::::
Db    1385970 AAAGTGATTGATGAG---------TCGCTTTTTAAAGAAATAAACGAGCGAAATATCGAA 1386020

Qy         28 ThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgPro 47
              :::|||:::|||||||||||||||||||:::   |||||||||||||||||||||||
Db    1386021 CCTAGACTAAAAGAGTCCATGTTATATTCTATTCAAGCAGGCGGGAAACGGATTCGTCCA 1386080

Qy         48 ValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSer 67
              :::|||:::      ||||||   :::|||              ||||||:::|||:::
Db    1386081 ATGCTAGTTTTCGCTACGCTTCAAGCTTTAAAAGTAAATCCGCTTTTAGGTGTAAAAACC 1386140

Qy         68 AlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMet 87
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1386141 GCAACAGCGTTAGAAATGATTCACACATACAGCCTTATTCATGATGATCTACCAGCAATG 1386200

Qy         88 AspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrpThr 107
              ||||||||||||||||||||||||||   ||||||||||||||:::|||:::   |||
Db    1386201 GATAACGATGATTATCGTCGAGGCAAGTGGACTAATCATAAAGTTTTTGGCGATGCAACT 1386260

Qy        108 AlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAsp 127
              ||||||||||||||||||||||||||||||   ||||||   :::::::::   ||||||

Db    1386261 GCGATTTTGGCAGGAGATGCTTTACTAACGCTCGCTTTTTCTATTTTAGCTGAAGACGAT 1386320

Qy        128 ArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHis 147
              |||:::   |||   ::::|||   ::::::   :::::::|||   :::||||||
Db    1386321 AATTTATCTTTTGAGACACGCATTGCTTTGATTAACCAAATTAGTTTTAGTAGCGGTGCA 1386380

Qy        148 ValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIleAspLeu 167
              ||||||||||||||||||:::   |||:::::::::|||   :::   :::   |||
Db    1386381 GAAGGAATGGTTGGTGGTCAACTTGCAGACTTGGAAGCGGAAAACAAACAAGTGACGCTA 1386440

Qy        168 GluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMet 187
              |||   |||       ||||||       ||||||||   ||||||   :::||||||
Db    1386441 GAAGAGTTATCATCCATTCATGCACGAAAAACGGGTGAATTATTAATTTATGCTGTAACC 1386500

Qy        188 SerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyrSer 207
              |||||||||   ||||||      |||          :::   |||      ::::::
Db    1386501 TCTGCAGCAAAAATTGCGGAAGCTGATCCAGAACAAACGAAACGCTTACGAATTTTTGCA 1386560

Qy        208 TyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGlyAspGlu 227
              ::::::|||:::   |||||||||   ||||||:::||||||      |||||||||
Db    1386561 GAGAATATTGGGATTGGATTTCAAATTAGCGACGATATTTTAGATGTAATTGGTGATGAA 1386620

Qy        228 AlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyrValSer 247
              |||:::|||||||||   |||:::|||          ||||||||||||||
Db    1386621 ACGAAAATGGGTAAAAAGACAGGGGCCGACGCTTTTCTGAATAAAAGTACCTATCCCGGA 1386680

Qy        248 LeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAlaValAsp 267
              ||||||      |||||||||:::      |||      |||      |||
Db    1386681 TTACTCACGCTTGATGGGGCAAAAAGGGCATTAAATGAGCATGTTACGATTGCAAAGTCA 1386740

Qy        268 GluLeuThrGlnIleAspGluGlnPheAsnThrLysHisLeuLeuGluIleValAspLeu 287
              |||:::      |||      |||:::   :::   ||||||:::::   ||||||
Db    1386741 GCGCTTTCAGGGCATGAT------TTCGATGATGAAATTCTCTTGAAACTTGCTGATTTA 1386794

RESULT 11
ABQ71010
ID   ABQ71010 standard; DNA; 5635 BP.
XX
AC   ABQ71010;
XX
DT   29-AUG-2002 (first entry)
XX
DE   Listeria monocytogenes 4b contig DNA sequence #952.
XX
KW   Antibacterial; Listeria; food contamination; mutational analysis;
KW   infection; ds.
XX
OS   Listeria monocytogenes 4b.
XX
PN   WO200228891-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-FR03061.
XX
PR   04-OCT-2000; 2000FR-0012697.
XX
PA   (INSP ) INST PASTEUR.
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Kunst F, Glaser P;
XX
DR   WPI; 2002-332479/37.
XX
PT   New genomic sequences from Listeria species, useful for detection,
PT   treatment and prevention of infection, also related polypeptides,
PT   antibodies and modulators -
XX
PS   Claim 14; SEQ ID 3823; 180pp; French.
XX
CC   The present invention relates to nucleic acid sequences
CC   (ABQ67188-ABQ71212) from Listeria sp. The sequences are useful as probes
CC   and primers for identification and/or detection of Listeria (e.g. as
```

```
CC   contaminants in foods, or mutational analysis) and for analysis of
CC   gene expression. Proteins encoded by the nucleic acid sequences can be
CC   used to screen for compounds that modulate gene expression, replication
CC   and pathogenicity of Listeria (potential therapeutic agents), also for
CC   treating infections by Listeria, and are useful as immunogens in
CC   anti-Listeria vaccines.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 5635 BP; 1896 A; 927 C; 1238 G; 1574 T; 0 other;

Alignment Scores:
Pred. No.:              1.22e-64      Length:        5635
Score:                  695.50        Matches:       148
Percent Similarity:     68.93%        Conservative:  45
Best Local Similarity:  52.86%        Mismatches:    82
Query Match:            47.87%        Indels:        5
DB:                     24            Gaps:          2

-09-925-637-64 (1-287) x ABQ71010 (1-5635)

Qy          8 LysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerValMetAsp 27
              |||:::|||||||||          |||        |||::::::      ::::::
Db       2914 AAAGTGATTGATGAG---------TCGCTTTTTAAAGAAATAAGCGAGCGAAATATCGAA 2964

Qy         28 ThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgPro 47
                ::::|||:::|||||||||||||||||:::   ||||||||||||||||||||||
Db       2965 CCTAGACTAAAAGAGTCCATGTTATATTCTGTTCAAGCAGGTGGAAAACGGATTCGTCCC 3024

Qy         48 ValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSer 67
              :::|||:::      ||||||  :::|||                |||||:::|||:::
Db       3025 ATGCTAGTTTTTGCTACACTTCAAGCCTTAAAAGTAAATCCGCTTTTAGGTGTGAAAACA 3084

Qy         68 AlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMet 87
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3085 GCAACGGCGTTAGAAATGATTCATACGTATAGCCTGATCCATGATGATTTACCAGCAATG 3144

Qy         88 AspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrpThr 107
              ||||||||||||||||||||||||   ||||||||||||||:::|||:::   |||
Db       3145 GATAATGATGACTATCGTCGCGGCAAATATACTAATCATAAAGTTTTTGGGGATGCAACT 3204

Qy        108 AlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAsp 127
              ||||||||||||||||||||||||||||   ||||||   :::::::::   |||:::
Db       3205 GCGATTTTGGCAGGAGACGCTTTACTAACGCTTGCTTTTTCTATTTTAGCGGAAGACGAA 3264

Qy        128 ArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHis 147
              |||:::   |||   :::|||   ::::::   :::::::|||   :::||||||
          3265 AATTTATCTTTTGAAACACGCATAGCTTTGATTAATCAAATTAGTTTTAGTAGTGGTGCA 3324

             148 ValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIleAspLeu 167
              ||||||||||||||||||:::   ||||||::::::|||   :::   :::   |||
Db       3325 GAAGGAATGGTTGGCGGTCAACTTGCAGATATGGAGGCAGAAAACAAACAAGTGACATTA 3384

Qy        168 GluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMet 187
              |||   |||      ||||||     |||||||||   ||||||   |||||||||
Db       3385 GAAGAACTATCATCCATCCATGCACGGAAAACAGGCGAATTATTAATTTTTGCCGTAACT 3444

Qy        188 SerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyrSer 207
              |||||||||   ||||||      |||          :::   |||      ::::::
Db       3445 TCTGCTGCGAAAATTGCAGAAGCTGACCCGGAACAAACGAAAAGATTACGGATTTTTGCG 3504

Qy        208 TyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGlyAspGlu 227
              ::::::|||:::   |||||||||   ||||||:::|||||      |||||||||
Db       3505 GAAAATATTGGGATTGGATTCCAAATCAGCGACGATATTTTAGATGTTATTGGTGATGAA 3564

Qy        228 AlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyrValSer 247
              |||:::|||||||||   |||   |||          ||||||||||||||
Db       3565 ACAAAAATGGGTAAAAAGACAGGGGTCGACGCTTTTCTGAACAAAAGTACCTATCCCGGA 3624

Qy        248 LeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAlaValAsp 267
              ||||||      |||||||||:::      |||      |||          |||
Db       3625 TTACTCACGCTTGATGGTGCCAAAAGGGCATTAAATGAGCATGTTGCGATTGCAAAGTCA 3684

Qy        268 GluLeuThrGlnIleAspGluGlnPheAsnThrLysHisLeuLeuGluIleValAspLeu 287
                 |||:::      |||      |||:::   :::   ||||||::::::   ||||||
Db       3685 GCGCTTTCAGGGCATGAT------TTTGATGATGAAATTCTCTTGAAACTTGCTGATTTA 3738

RESULT 12
AAS50245/c
ID   AAS50245 standard; DNA; 413 BP.
XX
AC   AAS50245;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus cellular proliferation inhibitory sequence #1469.
XX
KW   Antisense; ss; prokaryotic cellular proliferation;
KW   antibiotic; antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Claim 1; Seq ID No 2822; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence is an antisense
CC   oligonucleotide of the invention.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 413 BP; 115 A; 89 C; 58 G; 151 T; 0 other;

Alignment Scores:
Pred. No.:              1.57e-65      Length:        413
Score:                  690.00        Matches:       137
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
```

```
Query Match:            47.49%      Indels:        0
DB:                     23          Gaps:          0

US-09-925-637-64 (1-287) x AAS50245 (1-413)

Qy       54 LeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMet 73
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      413 TTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATG 354

Qy       74 IleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArg 93
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      353 ATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGA 294

Qy       94 ArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAsp 113
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      293 CGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGAT 234

Qy      114 AlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAspArgLeuThrAspGluVal 133
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      233 GCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTA 174

Qy      134 LysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGly 153
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      173 AAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGT 114

Qy      154 GlnMetLeuAspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIle 173
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      113 CAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATA 54

Qy      174 HisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAla 190
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       53 CACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCA 3

RESULT 13
ABQ67195/c
ID   ABQ67195 standard; DNA; 495269 BP.
XX
AC   ABQ67195;
XX
DT   29-AUG-2002  (first entry)
XX
DE   Listeria innocua contig DNA sequence #8.
XX
KW   Antibacterial; Listeria; food contamination; mutational analysis;
KW   infection; ds.
XX
OS   Listeria innocua.

     WO200228891-A2.

PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-FR03061.
XX
PR   04-OCT-2000; 2000FR-0012697.
XX
PA   (INSP ) INST PASTEUR.
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Kunst F,  Glaser P;
XX
DR   WPI; 2002-332479/37.
XX
PT   New genomic sequences from Listeria species, useful for detection,
PT   treatment and prevention of infection, also related polypeptides,
PT   antibodies and modulators  -
XX
PS   Claim 5; SEQ ID 8; 180pp; French.
XX
CC   The present invention relates to nucleic acid sequences
CC   (ABQ67188-ABQ71212) from Listeria sp. The sequences are useful as probes
CC   and primers for identification and/or detection of Listeria (e.g. as
CC   contaminants in foods, or mutational analysis) and for analysis of
CC   gene expression. Proteins encoded by the nucleic acid sequences can be
CC   used to screen for compounds that modulate gene expression, replication
CC   and pathogenicity of Listeria (potential therapeutic agents), also for
CC   treating infections by Listeria, and are useful as immunogens in
CC   anti-Listeria vaccines.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 495269 BP; 159529 A; 88849 C; 93135 G; 153754 T; 2 other;

Alignment Scores:
Pred. No.:              3.94e-61    Length:        495269
Score:                  686.50      Matches:       146
Percent Similarity:     68.21%      Conservative:  45
Best Local Similarity:  52.14%      Mismatches:    84
Query Match:            47.25%      Indels:        5
DB:                     24          Gaps:          2

US-09-925-637-64 (1-287) x ABQ67195 (1-495269)

Qy        8 LysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerValMetAsp 27
            |||:::::|||||           |||         |||||         ::::::
Db   366194 AAAGTACTTGATGAG---------TCGCTTTTTAAAGAAATAAATATGCGCAATATCGAA 366144

Qy       28 ThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgPro 47
              :::|||:::|||||||||||||||||||::::   |||||||||||||||||||||||
Db   366143 CCTAAGTTGAAAGAATCAATGTTATATTCAGTGCAAGCTGGCGGAAAACGAATTCGTCCA 366084

Qy       48 ValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSer 67
            :::|||:::      ||||||   :::|||||   |||       |||:::|||:::
Db   366083 ATGCTTGTTTTTGCAACGCTTCAAGCCTTAAATATTGAGCCAATGCGCGGTTTAAAAACA 366024

Qy       68 AlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMet 87
            |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   366023 GCTACGGCACTTGAAATGATTCATACGTACAGCTTAATTCACGATGATTTACCAGCAATG 365964

Qy       88 AspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrpThr 107
            |||:::|||||||||||||||||||||   ||||||||||||:::||||||:::   |||
Db   365963 GATGATGATGACTATCGACGCGGGAAATGGACGAACCATAAAATATACGGTGATGCAACA 365904

Qy      108 AlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAsp 127
            ||||||||||||||||||||||||||||||||   ||||||   :::::::::   |||:::
Db   365903 GCAATTTTAGCAGGAGATGCTTTGTTAACACTCGCTTTTTCTATTTTAGCTGAAGATGAA 365844

Qy      128 ArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHis 147
            |||:::   |||   :::|||   ::::::   :::::::|||   :::||||||
Db   365843 AATCTATCTTTCGAAACGCGTATTGCTTTAATTAATCAAATTAGTTATAGCAGTGGAGCA 365784

Qy      148 ValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIleAspLeu 167
            ||||||||||||||||||      ||||||::::::|||   :::   :::   |||
Db   365783 GAAGGCATGGTAGGTGGGCAACAAGCCGATATGGAAGCTGAAAATAAACAAGTCACATTA 365724

Qy      168 GluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMet 187
            |||   |||      ||||||      |||||||||   ||||||   |||||||||
Db   365723 GAAGAACTAGCATCAATCCACGCTCGCAAAACTGGTGAACTATTAATTTTCGCAGTAACC 365664

Qy      188 SerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyrSer 207
            |||||||||   ||||||                  :::   |||      ::::::
Db   365663 TCAGCCGCAAAAATCGCTGAAGCAACTCCAGAACAAACAAAACGATTACGAATTTTTGCA 365604

Qy      208 TyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGlyAspGlu 227
            ::::::|||:::   |||||||||   ||||||:::|||||||      |||||||||
Db   365603 GAAAATATCGGCATTGGTTTTCAAATTAGCGACGATATTTTAGATGTAATTGGCGATGAA 365544

Qy      228 AlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyrValSer 247
            |||:::||||||||||   |||   |||          ||||||||||||||||
Db   365543 ACAAAAATGGGTAAAAAGACAGGGGTCGACGCTTTTCTGAATAAAAGTACCTATCCCGGA 365484

Qy      248 LeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAlaValAsp 267
            ||||||      :::||||||:::      |||      |||      |||
Db   365483 TTACTCACACTCGAAGGTGCTAAACGGGCATTAAATGAGCATGTTTCAATTGCAAAGTCA 365424
```

```
Qy        268 GluLeuThrGlnIleAspGluGlnPheAsnThrLysHisLeuLeuGluIleValAspLeu 287
              |||:::      |||      |||:::   :::   ||||||:::::::   ||||||
Db     365423 GCGCTTTCAGGACATGAT-----TTCGACGATGAAATTCTCCTTAAACTTGCTGATTTA 365370

RESULT 14
ABQ69245
ID   ABQ69245 standard; DNA; 3011208 BP.
XX
AC   ABQ69245;
XX
DT   29-AUG-2002  (first entry)
XX
DE   Listeria innocua DNA sequence #684.
XX
KW   Antibacterial; Listeria; food contamination; mutational analysis;
KW   infection; ds.
XX
OS   Listeria innocua.
XX
PN   WO200228891-A2.
XX
     11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-FR03061.
XX
PR   04-OCT-2000; 2000FR-0012697.
XX
PA   (INSP ) INST PASTEUR.
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Kunst F,  Glaser P;
XX
DR   WPI; 2002-332479/37.
XX
PT   New genomic sequences from Listeria species, useful for detection,
PT   treatment and prevention of infection, also related polypeptides,
PT   antibodies and modulators -
XX
PS   Claim 5; SEQ ID 2058; 180pp; French.
XX
CC   The present invention relates to nucleic acid sequences
CC   (ABQ67188-ABQ71212) from Listeria sp. The sequences are useful as probes
CC   and primers for identification and/or detection of Listeria (e.g. as
CC   contaminants in foods, or mutational analysis) and for analysis of
CC   gene expression. Proteins encoded by the nucleic acid sequences can be
CC   used to screen for compounds that modulate gene expression, replication
CC   and pathogenicity of Listeria (potential therapeutic agents), also for
CC   treating infections by Listeria, and are useful as immunogens in
     anti-Listeria vaccines.
     Note: The sequence data for this patent did not form part
     of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 3011208 BP; 941651 A; 568176 C; 559189 G; 942192 T; 0 other;

Alignment Scores:
Pred. No.:              4.16e-60      Length:         3011208
Score:                  686.50        Matches:        146
Percent Similarity:     68.21%        Conservative:   45
Best Local Similarity:  52.14%        Mismatches:     84
Query Match:            47.25%        Indels:         5
DB:                     24            Gaps:           2

US-09-925-637-64 (1-287) x ABQ69245 (1-3011208)

Qy          8 LysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerValMetAsp 27
              |||::::::||||||          |||        ||||||        ::::::
Db    1392757 AAAGTACTTGATGAG---------TCGCTTTTTAAAGAAATAAATATGCGCAATATCGAA 1392807

Qy         28 ThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgPro 47
                ::::|||::::|||||||||||||||||:::   ||||||||||||||||||||||||
Db    1392808 CCTAAGTTGAAAGAATCAATGTTATATTCAGTGCAAGCTGGCGGAAAACGAATTCGTCCA 1392867

Qy         48 ValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSer 67
              :::|||:::      ||||||   :::|||||   |||        |||:::|||:::
Db    1392868 ATGCTTGTTTTTGCAACGCTTCAAGCCTTAAATATTGAGCCAATGCGCGGTTTAAAAACA 1392927

Qy         68 AlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMet 87
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1392928 GCTACGGCACTTGAAATGATTCATACGTACAGCTTAATTCACGATGATTTACCAGCAATG 1392987

Qy         88 AspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrpThr 107
              |||:::||||||||||||||||||||   ||||||||||||:::||||||:::   |||
Db    1392988 GATGATGATGACTATCGACGCGGGAAATGGACGAACCATAAAATATACGGTGATGCAACA 1393047

Qy        108 AlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAsp 127
              ||||||||||||||||||||||||||||||   ||||||   :::::::::   |||:::
Db    1393048 GCAATTTTAGCAGGAGATGCTTTGTTAACACTCGCTTTTTCTATTTTAGCTGAAGATGAA 1393107

Qy        128 ArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHis 147
              |||:::   |||   :::|||   ::::::   :::::::|||   :::||||||
Db    1393108 AATCTATCTTTCGAAACGCGTATTGCTTTAATTAATCAAATTAGTTATAGCAGTGGAGCA 1393167

Qy        148 ValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIleAspLeu 167
              ||||||||||||||||||      ||||||::::::|||   :::   :::   |||
Db    1393168 GAAGGCATGGTAGGTGGGCAACAAGCCGATATGGAAGCTGAAAATAAACAAGTCACATTA 1393227

Qy        168 GluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMet 187
              |||   |||      ||||||      |||||||||   ||||||   |||||||||
Db    1393228 GAAGAACTAGCATCAATCCACGCTCGCAAAACTGGTGAACTATTAATTTTCGCAGTAACC 1393287

Qy        188 SerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyrSer 207
              |||||||||   ||||||                        :::   |||      ::::::
Db    1393288 TCAGCCGCAAAAATCGCTGAAGCAACTCCAGAACAAACAAAACGATTACGAATTTTTGCA 1393347

Qy        208 TyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGlyAspGlu 227
              ::::::|||:::   |||||||||   ||||||:::||||||     |||||||||
Db    1393348 GAAAATATCGGCATTGGTTTTCAAATTAGCGACGATATTTTAGATGTAATTGGCGATGAA 1393407

Qy        228 AlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyrValSer 247
              |||:::||||||||||   |||   |||        ||||||||||||||||
Db    1393408 ACAAAAATGGGTAAAAAGACAGGGGTCGACGCTTTTCTGAATAAAAGTACCTATCCCGGA 1393467

Qy        248 LeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAlaValAsp 267
              ||||||      :::||||||:::      |||      |||          |||
Db    1393468 TTACTCACACTCGAAGGTGCTAAACGGGCATTAAATGAGCATGTTTCAATTGCAAAGTCA 1393527

Qy        268 GluLeuThrGlnIleAspGluGlnPheAsnThrLysHisLeuLeuGluIleValAspLeu 287
              |||:::      |||      |||:::   :::   ||||||:::::::   ||||||
Db    1393528 GCGCTTTCAGGACATGAT------TTCGACGATGAAATTCTCCTTAAACTTGCTGATTTA 1393581

RESULT 15
AAT40228
ID   AAT40228 standard; DNA; 894 BP.
XX
AC   AAT40228;
XX
DT   11-MAY-1997  (first entry)
XX
DE   Mutant farnesyldiphosphate synthase (4).
XX
KW   Farnesyldiphosphate; FPP; synthase; mutant; enzyme;
KW   geranylgeranyl diphosphate; GGPP; ds.
XX
OS   Bacillus stearothermophilus.
XX
PN   EP733709-A2.
XX
PD   25-SEP-1996.
XX
PF   29-SEP-1995;   95EP-0115423.
XX
PR   14-FEB-1995;   95JP-0025253.
XX
PA   (TOYT ) TOYOTA JIDOSHA KK.
XX
```

PI Ayumi K, Kyozo O, Shinichi O, Shusei O, Takeshi N;
PI Tanetoshi K, Tokuzo N;
XX
DR WPI; 1996-427057/43.
DR P-PSDB; AAW00285.
XX
PT Production of geranyl:geranyl di:phosphate synthase enzymes - by
PT mutation of farnesyl di:phosphate synthase
XX
PS Claim 18; Page 15-16; 50pp; English.
XX
CC Four mutant FPP synthase genes are given in AAT40225 to AAT40228. The
CC native FPP synthase gene is given in AAT40229. The mutants are
CC capable of synthesising geranylgeranyl diphosphate (GGPP) or
CC geranylgeranyol from isopentenyl, dimethylallyl, geranyl or farnesyl
CC diphosphate. The GGPP synthase is produced from FPP synthase by
CC deletion, addition or replacement of one to a few amino acids. The
CC modification is present in at least one of the following positions:
CC 34, 59, 81, 157, 182, 239, 265 and 275.

Sequence 894 BP; 207 A; 243 C; 274 G; 170 T; 0 other;

Alignment Scores:

Pred. No.:	1.57e-61	Length:	894
Score:	657.00	Matches:	133
Percent Similarity:	68.10%	Conservative:	57
Best Local Similarity:	47.67%	Mismatches:	83
Query Match:	45.22%	Indels:	6
DB:	17	Gaps:	2

US-09-925-637-64 (1-287) x AAT40228 (1-894)

```
Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              |||      |||   :::   :::   ::::::|||           :::      |||
Db          1 ATGGCGCAGCTTTCAGTTGAACAGTTTCTCAACGAGCAAAAACAGGCGGTGGAAACAGCG 60

Qy         21 IleAsnLysSerValMet----------AspThrGlnLeuGluGluSerMetLeuTyr 36
              :::::::::   :::                :::|||::::::::::|||   |||
Db         61 CTCTCCCGTTATATAGAGCGCTTAGAAGGGCCGGCGAAGCTGAAAAAGGCGATGGCGTAC 120

Qy         37 SerLeuAsnAlaGlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSer 56
              ||||||   ||||||||||||||||||||||||:::|||||||||   |||:::   :::
Db        121 TCATTGGAGGCCGGCGGCAAACGAATCCGTCCGTTGCTGCTTCTGTCCACCGTTCGGGCG 180

Qy         57 LeuAsnThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThr 76
              |||      :::      :::|||:::      |||   |||:::|||||||||||||||
Db        181 CTCGGCAAAGACCCGGCGGTCGGATTGCCCGTCGCCTGCGCGATTGAAATGATCCATACG 240

Qy         77 TyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLys 96
              :::||||||||||||||||||||||||:::|||||||||||||   ||||||||||||
Db        241 CACTCTTTGATCCATGATGATTTGCCGAGCATGGACAACGATGATTTGCGGCGCGGCAAG 300

Qy         97 LeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeu 116
                ||||||||||||||:::|||||||      |||||||||||||||||   ||||||
Db        301 CCGACGAACCATAAAGTGTTCGGCGAGGCGATGGCCATCTTGGCGGGGGACGGGTTGTTG 360

Qy        117 ThrLysAlaPheGluLeuIleSer------SerAspAspArgLeuThrAspGluValLys 134
              |||   ||||||:::||||||:::         |||:::|||:::         |||:::
Db        361 ACGTACGCGTTTCAATTGATCACCGAAATCGACGATGAGCGCATCCCTCCTTCCGTCCGG 420

Qy        135 IleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGln 154
              ::::::::::::::|||||:::   |||:::|||      |||||||||   ||||||
Db        421 CTTCGGCTCATCGAACGGCTGGCGAAAGCGGCCGGTCCGGAAGGGATGGTCGCCGGTCAG 480

Qy        155 MetLeuAspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHis 174
                  ||||||:::   ||||||:::   :::   |||      ||||||   ||||||
Db        481 GCAGCCGATATGGAAGGAGAGGGGAAAACGCTGACGCTTTCGGAGCTCGAATACATTCAT 540

Qy        175 LysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsn 194
              :::   |||||||||   :::|||   ::::::|||   :::   |||   |||
Db        541 CGGCATAAAACCGGGAAAATGCTGCAATACAGCGTGCACGCCGGCGCCTTGATCGGCGGC 600

Qy        195 ValAspAspThrThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPhe 214
                 |||                  |||:::   ::::::   ||||||||||:::   |||
Db        601 GCTGATGCCCGGCAAACGCGGGAGCTTGACGAATTCGCCGCCCATCTAGGCCTTGCCTTT 660

Qy        215 GlnIleLysAspAspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysVal 234
              ||||||:::|||||||:::||||||      |||   |||   |||:::|||||||:::|||
Db        661 CAAATTCGCGATGATATTCTCGATATTGAAGGGGCAGAAGAAAAAATCGGCAAGCGGGTC 720

Qy        235 GlySerAspLeuGluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAla 254
              |||||||||      |||||||||:::||||||   :::||||||         ||||||
Db        721 GGCAGCGACCAAAGCAACAACAAAGCGACGTATCCAGCGTTGCTGTCGCTTGCCGGCGCG 780

Qy        255 GluAspLysLeuThrTyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAsp 273
              ::::::||||||||||:::|||   :::||||||         |||         |||
Db        781 AAGGAAAAGTTGACGTTCCATATCGAGGCGGCGCAGCGCCATTTACGGAACGCCGAC 837
```

Search completed: May 30, 2003, 20:55:05
Job time : 1358 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 30, 2003, 20:32:16 ; Search time 59 Seconds
(without alignments)
1491.801 Million cell updates/sec

Title: US-09-925-637-64
Perfect score: 1453
Sequence: 1 MTNLPMNKLIDEVNNELSVA.....ELTQIDEQFNTKHLLEIVDL 287

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

arched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09925637/runat_23052003_174818_23446/app_query.fasta_1.455
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09925637_@CGN_1_1_40_@runat_23052003_174818_23446 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120. -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1436	98.8	882	3	US-09-276-873-1	Sequence 1, Appli
2	1015	69.9	909	4	US-09-134-001C-334	Sequence 334, App
3	657	45.2	894	1	US-08-534-910B-4	Sequence 4, Appli
4	656	45.1	894	1	US-08-534-910B-5	Sequence 5, Appli
5	654	45.0	893	1	US-08-333-321-1	Sequence 1, Appli
6	653	44.9	894	1	US-08-534-910B-2	Sequence 2, Appli
7	650	44.7	894	1	US-08-534-910B-3	Sequence 3, Appli
8	650	44.7	894	3	US-08-886-466-1	Sequence 1, Appli
9	650	44.7	894	4	US-09-475-304-1	Sequence 1, Appli
10	650	44.7	894	4	US-09-101-126-4	Sequence 4, Appli
11	650	44.7	894	4	US-09-367-528A-4	Sequence 4, Appli
12	645	44.4	894	1	US-08-534-910B-1	Sequence 1, Appli
13	644	44.3	894	4	US-09-367-528A-2	Sequence 2, Appli
14	551	37.9	876	3	US-09-275-742-1	Sequence 1, Appli
c 15	549	37.8	10011	4	US-08-961-527-76	Sequence 76, Appl
16	520.5	35.8	885	3	US-09-187-050-11	Sequence 11, Appl
17	520.5	35.8	1179	3	US-09-187-050-13	Sequence 13, Appl
18	520.5	35.8	1179	3	US-09-187-050-15	Sequence 15, Appl
19	520.5	35.8	1179	3	US-09-187-050-17	Sequence 17, Appl
20	520.5	35.8	1179	3	US-09-187-050-19	Sequence 19, Appl
21	520.5	35.8	1179	3	US-09-187-050-21	Sequence 21, Appl
22	520.5	35.8	1179	3	US-09-187-050-23	Sequence 23, Appl
23	520.5	35.8	1179	3	US-09-187-050-25	Sequence 25, Appl
24	520.5	35.8	1889	3	US-09-187-050-1	Sequence 1, Appli
25	465.5	32.0	1131	4	US-09-420-211-1	Sequence 1, Appli
26	438.5	30.2	486	3	US-09-217-609A-28	Sequence 28, Appl
27	438.5	30.2	486	4	US-08-873-235B-28	Sequence 28, Appl
28	376.5	25.9	909	1	US-07-783-705A-7	Sequence 7, Appli
29	376.5	25.9	6918	1	US-07-783-705A-13	Sequence 13, Appl
30	344	23.7	1284	4	US-08-858-207A-173	Sequence 173, App
31	340	23.4	1157	1	US-08-095-726-1	Sequence 1, Appli
32	340	23.4	1157	1	US-08-095-726-3	Sequence 3, Appli
33	340	23.4	1157	1	US-08-096-043-1	Sequence 1, Appli
34	340	23.4	1157	1	US-08-096-043-3	Sequence 3, Appli
35	340	23.4	1157	1	US-08-093-577-1	Sequence 1, Appli
36	340	23.4	1157	1	US-08-093-577-3	Sequence 3, Appli
37	340	23.4	1157	1	US-08-096-623A-1	Sequence 1, Appli
38	340	23.4	1157	1	US-08-096-623A-3	Sequence 3, Appli
39	332	22.8	993	1	US-08-705-377-2	Sequence 2, Appli
40	332	22.8	993	2	US-09-052-962-2	Sequence 2, Appli
41	332	22.8	993	2	US-09-053-068-2	Sequence 2, Appli
42	330	22.7	990	1	US-08-410-167A-1	Sequence 1, Appli
43	330	22.7	993	1	US-08-705-377-1	Sequence 1, Appli
44	330	22.7	993	2	US-09-052-962-1	Sequence 1, Appli
45	330	22.7	993	2	US-09-053-068-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-276-873-1
; Sequence 1, Application US/09276873
; Patent No. 6107058
; GENERAL INFORMATION:
; APPLICANT: Wilding, Edwina Imogen
; APPLICANT: Gwynn, Michael
; TITLE OF INVENTION: IspA
; FILE REFERENCE: GM10208
; CURRENT APPLICATION NUMBER: US/09/276,873
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 882
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-276-873-1

Alignment Scores:
Pred. No.: 4.34e-168 Length: 882
Score: 1436.00 Matches: 284
Percent Similarity: 99.30% Conservative: 1
Best Local Similarity: 98.95% Mismatches: 2
Query Match: 98.83% Indels: 0
DB: 3 Gaps: 0

US-09-925-637-64 (1-287) x US-09-276-873-1 (1-882)

```
Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGCATGTTGTATTCATTAAATGCT 120
```

```
Qy        41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy        61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
             ||||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 TATGAGTTAGGTGTGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy        81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
             ||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db       241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAAGGAAATTAACAAATCAT 300

Qy       101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy       121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy       141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy       161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy       181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys 200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||
Db       541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAA 600

Qy       201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660

Qy       221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Qy       241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy       261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy       281 LeuLeuGluIleValAspLeu 287
             |||||||||||||||||||||
Db       841 TTATTAGAAATCGTTGATTTA 861

RESULT 2
US-09-134-001C-334
; Sequence 334, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 334
;   LENGTH: 909
;   TYPE: DNA
;   ORGANISM: Staphylococcus epidermidis

US-09-134-001C-334

Alignment Scores:
Pred. No.:              4.36e-116      Length:          909
Score:                  1015.00        Matches:         197
Percent Similarity:     82.93%         Conservative:    41
Best Local Similarity:  68.64%         Mismatches:      49
Query Match:            69.86%         Indels:          0
DB:                     4              Gaps:            0

US-09-925-637-64 (1-287) x US-09-134-001C-334 (1-909)

Qy         1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
             |||      |||   |||||||||||||||:::   :::|||      |||:::   :::
Db        28 ATGAAGAAACTACAGATGAATAAATTAATAAATATAATTAATACATCACTGAATAAGTCA 87

Qy        21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
             |||      |||   :::   |||   ||||||||||||||   ||||||||||||||||
Db        88 ATACAATCATCACCATTAAAAACTAATTTAGAAGAAAGTATGAAATATTCATTAAATGCT 147

Qy        41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
             ||||||||||||||||||||||:::|||||||||||||||      ||||||   :::
Db       148 GGTGGTAAAAGAATCAGACCAGTCATATTATTATTAACACTAAAAATGCTTAACAAAGAT 207

Qy        61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
             |||:::   |||:::   |||||||:::||||||||||||||||||||||||||||||||
Db       208 TATCAACAAGGACTAAATAGTGCTTTAGCATTGGAAATGATTCATACTTATTCTTTAATT 267

Qy        81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       268 CATGATGATTTACCAGCAATGGATAATGACGATTACCGTAGAGGAAAATTAACAAATCAT 327

Qy       101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
             ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db       328 AAAGTTTATGGTGAATGGAAAGCCATTCTTGCTGGTGATGCATTATTAACAAAAGCTTTT 387

Qy       121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
             ||||||:::|||:::|||      :::   |||   ||||||:::   :::::::::|||
Db       388 GAATTAGTTTCTAATGATACTACCATTGAAGATAGTGTGAAAGTAAGTATTATAAAAAGA 447

Qy       141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
             ||||||   ||||||||||||:::||||||||||||||||   ||||||||||:::|||
Db       448 CTTTCAAAAGCAAGTGGACATTTGGGAATGGTGGGTGGCCAAGCGCTTGATATGGAAAGT 507

Qy       161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
             ||||||:::   |||   ||||||||||||||   ||||||:::||||||||||||||||
Db       508 GAAGGGAAGTCAATTCGTTTAGAAACTTTAGAATCAATTCATGAAACTAAGACAGGCGCT 567

Qy       181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys 200
             ||||||   |||:::||||||:::|||   |||||||||   |||:::
Db       568 TTACTAAATTTTTCAGTTATGGCTGCGGTAGACATTGCTCAAGTAGAACAAAATATTGCT 627

Qy       201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
             ::::::|||:::   :::|||:::||||||||||||||||||||||||||||||||||||
Db       628 AAGAATTTAGATGAATTTAGTCATCATTTAGGAATGATGTTTCAAATTAAAGATGATTTA 687

Qy       221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
             ||||||   ||||||||||||:::||||||||||||||||||||||||||:::   |||
Db       688 CTGGATGTGTATGGTGATGAATCAAAACTTGGCAAAAAAGTAGGCAGTGATATAGTAAAT 747

Qy       241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
             :::||||||||||||||||||||||||||||||:::||||||||||:::||||||
Db       748 CATAAAAGTACTTATGTTTCTTTACTTGGAAAAGAAGGAGCAGAAGAAAAGTTAAACAAT 807

Qy       261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
             |||:::      |||::::::   |||   ||||||   :::|||::::::|||
Db       808 CATCAATATCTTGCTATGAACTGCTTAAATCAAATTTCTGATCAATATGATACTTCTGAA 867

Qy       281 LeuLeuGluIleValAspLeu 287
             |||   :::||||||||||||
Db       868 TTAAGTGATATTGTAGATTTA 888

RESULT 3
```

```
US-08-534-910B-4
; Sequence 4, Application US/08534910B
; Patent No. 5766911
;  GENERAL INFORMATION:
;    APPLICANT:  KOIKE, Ayumi
;    APPLICANT:  OBATA, Shusei
;    APPLICANT:  NISHINO, Tokuzo
;    APPLICANT:  OHNUMA, Shinichi
;    APPLICANT:  NAKAZAWA, Takeshi
;    APPLICANT:  OGURA, Kyozo
;    APPLICANT:  KOYAMA, Tanetoshi
;    TITLE OF INVENTION:  Mutated Farnesyldiphosphate Synthase Capable
;    TITLE OF INVENTION:  Of Synthesizing Geranylgeranyldiphosphate And Gene Coding Ther
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenyon & Kenyon
;      STREET:  1025 Connecticut Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  U.S.
;      ZIP:  20036-5405
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.25" Floppy Disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;      SOFTWARE:  IBM/Word Perfect 6.1 Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/534,910B
;      FILING DATE:  28-SEPT-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-25253
;      FILING DATE:  14-FEB-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Toffenetti, Judith L.
;      REGISTRATION NUMBER:  39,048
;      REFERENCE/DOCKET NUMBER:  77670/398
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)429-1776
;      TELEFAX:  (202)429-0796
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  894 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Bacillus stearothermophilus
   08-534-910B-4

Alignment Scores:
Pred. No.:              6.79e-72        Length:         894
Score:                  657.00          Matches:        133
Percent Similarity:     68.10%          Conservative:   57
Best Local Similarity:  47.67%          Mismatches:     83
Query Match:            45.22%          Indels:         6
DB:                     1               Gaps:           2

US-09-925-637-64 (1-287) x US-08-534-910B-4 (1-894)

Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              |||      |||   :::   :::   ::::::|||          ::::      |||
Db          1 ATGGCGCAGCTTTCAGTTGAACAGTTTCTCAACGAGCAAAAACAGGCGGTGGAAACAGCG 60

Qy         21 IleAsnLysSerValMet-----------AspThrGlnLeuGluGluSerMetLeuTyr 36
              :::::::::   :::                    :::|||:::::::::|||   |||
Db         61 CTCTCCCGTTATATAGAGCGCTTAGAAGGGCCGGCGAAGCTGAAAAAGGCGATGGCGTAC 120

Qy         37 SerLeuAsnAlaGlyGlyLysArgIleArgProValLeuLeuLeuThrLeuAspSer 56
              ||||||   ||||||||||||||||||||||||:::|||||||||   |||:::   :::
Db        121 TCATTGGAGGCCGGCGGCAAACGAATCCGTCCGTTGCTGCTTCTGTCCACCGTTCGGGCG 180

Qy         57 LeuAsnThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThr 76
              |||      :::      :::|||:::      |||   |||:::|||||||||||||||
Db        181 CTCGGCAAAGACCCGGCGGTCGGATTGCCCGTCGCCTGCGCGATTGAAATGATCCATACG 240

Qy         77 TyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLys 96
              :::|||||||||||||||||||||||||:::|||||||||||||||   ||||||||||
Db        241 CACTCTTTGATCCATGATGATTTGCCGAGCATGGACAACGATGATTTGCGGCGCGGCAAG 300

Qy         97 LeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeu 116
              ||||||||||||||||:::||||||      ||||||||||||||||||   ||||||
Db        301 CCGACGAACCATAAAGTGTTCGGCGAGGCGATGGCCATCTTGGCGGGGGACGGGTTGTTG 360

Qy        117 ThrLysAlaPheGluLeuIleSer------SerAspAspArgLeuThrAspGluValLys 134
              |||   ||||||:::||||||:::        |||:::|||:::        |||:::
Db        361 ACGTACGCGTTTCAATTGATCACCGAAATCGACGATGAGCGCATCCCTCCTTCCGTCCGG 420

Qy        135 IleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGln 154
              :::::::::::::::|||||:::   |||:::|||      ||||||||   ||||||
Db        421 CTTCGGCTCATCGAACGGCTGGCGAAAGCGGCCGGTCCGGAAGGGATGGTCGCCGGTCAG 480

Qy        155 MetLeuAspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHis 174
              ||||||:::   ||||||:::   :::   |||      ||||||   ||||||
Db        481 GCAGCCGATATGGAAGGAGAGGGGAAAACGCTGACGCTTTCGGAGCTCGAATACATTCAT 540

Qy        175 LysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsn 194
              :::   ||||||||   :::|||   ::::::|||   :::   |||   |||
Db        541 CGGCATAAAACCGGGAAAATGCTGCAATACAGCGTGCACGCCGGCGCCTTGATCGGCGGC 600

Qy        195 ValAspAspThrThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPhe 214
              |||                  |||:::   ::::::   |||||||||:::   |||
Db        601 GCTGATGCCCGGCAAACGCGGGAGCTTGACGAATTCGCCGCCCATCTAGGCCTTGCCTTT 660

Qy        215 GlnIleLysAspAspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysVal 234
              ||||||:::||||||:::||||||      |||   |||   |||:::||||||:::|||
Db        661 CAAATTCGCGATGATATTCTCGATATTGAAGGGGCAGAAGAAAAAATCGGCAAGCGGGTC 720

Qy        235 GlySerAspLeuGluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAla 254
              |||||||||      |||||||||:::||||||   :::|||||         ||||||
Db        721 GGCAGCGACCAAAGCAACAACAAAGCGACGTATCCAGCGTTGCTGTCGCTTGCCGGCGCG 780

Qy        255 GluAspLysLeuThrTyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAsp 273
              ::::::||||||||||:::|||   :::||||||         |||         |||
Db        781 AAGGAAAAGTTGACGTTCCATATCGAGGCGGCGCAGCGCCATTTACGGAACGCCGAC 837

RESULT 4
US-08-534-910B-5
; Sequence 5, Application US/08534910B
; Patent No. 5766911
;  GENERAL INFORMATION:
;    APPLICANT:  KOIKE, Ayumi
;    APPLICANT:  OBATA, Shusei
;    APPLICANT:  NISHINO, Tokuzo
;    APPLICANT:  OHNUMA, Shinichi
;    APPLICANT:  NAKAZAWA, Takeshi
;    APPLICANT:  OGURA, Kyozo
;    APPLICANT:  KOYAMA, Tanetoshi
;    TITLE OF INVENTION:  Mutated Farnesyldiphosphate Synthase Capable
;    TITLE OF INVENTION:  Of Synthesizing Geranylgeranyldiphosphate And Gene Coding T
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenyon & Kenyon
;      STREET:  1025 Connecticut Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  U.S.
;      ZIP:  20036-5405
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.25" Floppy Disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;      SOFTWARE:  IBM/Word Perfect 6.1 Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/534,910B
```

```
;      FILING DATE:  28-SEPT-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-25253
;      FILING DATE:  14-FEB-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Toffenetti, Judith L.
;      REGISTRATION NUMBER:  39,048
;      REFERENCE/DOCKET NUMBER:  77670/398
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)429-1776
;      TELEFAX:  (202)429-0796
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  894 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Bacillus stearothermophilus
US-08-534-910B-5

Alignment Scores:
Pred. No.:              9.03e-72      Length:          894
Score:                  656.00        Matches:         133
Percent Similarity:     67.38%        Conservative:    55
Best Local Similarity:  47.67%        Mismatches:      85
Query Match:            45.15%        Indels:          6
DB:                     1             Gaps:            2

US-09-925-637-64 (1-287) x US-08-534-910B-5 (1-894)

Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              |||      |||   :::   :::   ::::::|||          :::      |||
Db          1 ATGGCGCAGCTTTCAGTTGAACAGTTTCTCAACGAGCAAAAACAGGCGGTGGAAACAGCG 60

Qy         21 IleAsnLysSerValMet-----------AspThrGlnLeuGluGluSerMetLeuTyr 36
              :::::::::   :::                . :::|||::::::::::|||. |||
Db         61 CTCTCCCGTTATATAGAGCGCTTAGAAGGGCCGGCGAAGCTGAAAAAGGCGATGGCGTAC 120

Qy         37 SerLeuAsnAlaGlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSer 56
              ||||||   ||||||||||||||||||||||:::||||||||   |||:::   :::
Db        121 TCATTGGAGGCCGGCGGCAAACGAATCCGTCCGTTGCTGCTTCTGTCCACCGTTCAGGCG 180

Qy         57 LeuAsnThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThr 76
              |||      :::     :::|||:::      |||   |||:::|||||||||||||||
           181 CTCGGCAAAGACCCGGCGGTCGGATTGCCCGTCGCCTGCGCGATTGAAATGATCCATACG 240

            77 TyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLys 96
              ||||||||||||||||||||||||||:::|||||||||||||||   ||||||||||||
Db        241 TACTCTTTGATCCATGATGATTTGCCGAGCATGGACAACGATGATTTGCGGCGCGGCAAG 300

Qy         97 LeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeu 116
              ||||||||||||||:::||||||      ||||||||||||||||||   ||||||
Db        301 CCGACGAACCATAAAGTGTTCGGCGAGGCGATGGCCATCTTGGCGGGGGACGGGTTGTTG 360

Qy        117 ThrLysAlaPheGluLeuIleSer------SerAspAspArgLeuThrAspGluValLys 134
              |||   ||||||:::||||||:::        |||:::|||:::      |||:::
Db        361 ACGTACGCGTTTCAATTGATCACCGAAATCGACGATGAGCGCATCCCTCCTTCCGTCCGG 420

Qy        135 IleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGln 154
              ::::::::::::::::|||||||:::   |||:::|||      ||||||||   ||||||
Db        421 CTTCGGCTCATCGAACGGCTGGCGAAAGCGGCCGGTCCGGAAGGGATGGTCGCCGGTCAG 480

Qy        155 MetLeuAspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHis 174
                    ||||||:::   |||||||:::   :::   |||      ||||||   ||||||
Db        481 GCAGCCGATATGGAAGGAGAGGGGAAAACGCTGACGCTTTCGGAGCTCGAATACATTCAT 540

Qy        175 LysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsn 194
              :::   |||||||||   :::|||   ::::::|||   :::   |||   |||
Db        541 CGGCATAAAACCGGGAAAATGCTGCAATACAGCGTGCACGCCGGCGCCTTGATCGGCGGC 600

Qy        195 ValAspAspThrThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPhe 214
                 |||                ||||::   ::::::   |||||||||:::   |||
Db        601 GCTGATGCCCGGCAAACGCGGGAGCTTGACGAATTCGCCGCCCATCTAGGCCTTGCCTTT 660

Qy        215 GlnIleLysAspAspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysVal 234
              ||||||:::||||||:::||||||      |||   |||   |||:::||||||   |||
Db        661 CAAATTCGCGATGATATTCTCGATATTGAAGGGGCAGAAGAAAAAATCGGCAAGCCGGTC 720

Qy        235 GlySerAspLeuGluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAla 254
              |||||||||      ||||||||||:::||||||   :::||||||          ||||||
Db        721 GGCAGCGACCAAAGCAACAACAAAGCGACGTATCCAGCGTTGCTGTCGCTTGCCGGCGCG 780

Qy        255 GluAspLysLeuThrTyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAsp 273
              ::::::||||||   :::|||   :::||||||          |||          |||
Db        781 AAGGAAAAGTTGGCGTTCCATATCGAGGCGGCGCAGCGCCATTTACGGAACGCCGAC 837

RESULT 5
US-08-333-321-1
; Sequence 1, Application US/08333321
; Patent No. 5786192
;  GENERAL INFORMATION:
;    APPLICANT:  Obata, Shusei
;    TITLE OF INVENTION:  FARNESYL PYROPHOPHATE SYNTHETASE
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Edward W. Greason, Esq.
;      STREET:  1 Broadway
;      CITY:  New York
;      STATE:  N.Y.
;      COUNTRY:  U.S.A.
;      ZIP:  10004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  No. 5786192epad, Windows 3.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/333,321
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/953,424
;      FILING DATE:  29-SEP-1992
;      APPLICATION NUMBER:  JP 3-253788
;      FILING DATE:  01-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Greason, Edward W.
;      REGISTRATION NUMBER:  18,918
;      REFERENCE/DOCKET NUMBER:  077670/00310
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 425-7200
;      TELEFAX:  (212) 425-5288
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  893 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  Bacillus stearothermophilus
US-08-333-321-1

Alignment Scores:
Pred. No.:              1.59e-71      Length:          893
Score:                  654.00        Matches:         133
Percent Similarity:     67.38%        Conservative:    55
Best Local Similarity:  47.67%        Mismatches:      85
Query Match:            45.01%        Indels:          6
DB:                     1             Gaps:            2

US-09-925-637-64 (1-287) x US-08-333-321-1 (1-893)
```

```
Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              |||      |||   :::   :::   ::::::|||            :::      |||
Db          1 ATGGCGCAGCTTTCAGTTGAACAGTTTCTCAACGAGCAAAAACAGGCGGTGGAAACAGCG 60

Qy         21 IleAsnLysSerValMet----------AspThrGlnLeuGluGluSerMetLeuTyr 36
              :::::::::   :::                :::|||::::::::::|||   |||
Db         61 CTCTCCCGTTATATAGAGCGCTTAGAAGGGCCGGCGAAGCTGAAAAAGGCGATGGCGTAC 120

Qy         37 SerLeuAsnAlaGlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSer 56
              ||||||   ||||||||||||||||||||||:::|||||||||   |||:::   :::
Db        121 TCATTGGAGGCCGGCGGCAAACGAATCCGTCCGTTGCTGCTTCTGTCCACCGTTCGGGCG 180

Qy         57 LeuAsnThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThr 76
              |||      :::      :::|||:::      |||   |||:::|||||||||||||||
Db        181 CTCGGCAAAGACCCGGCGGTCGGATTGCCCGTCGCCTGCGCGATTGAAATGATCCATACG 240

Qy         77 TyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLys 96
              ||||||||||||||||||||||||||:::|||||||||||||||   ||||||||||||
Db        241 TACTCTTTGATCCATGATGATTTGCCGAGCATGGACAACGATGATTTGCGGCGCGGCAAG 300

Qy         97 LeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeu 116
              ||||||||||||||:::|||||||     ||||||||||||||||||   ||||||
Db        301 CCGACGAACCATAAAGTGTTCGGCGAGGCGATGGCCATCTTGGCGGGGGACGGGTTGTTG 360

Qy        117 ThrLysAlaPheGluLeuIleSer------SerAspAspArgLeuThrAspGluValLys 134
              |||   |||||:::|||||:::         |||:::|||:::         |||:::
Db        361 ACGTACGCGTTTCAATTGATCACCGAAATCGACGATGAGCGCATCCCTCCTTCCGTCCGG 420

Qy        135 IleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGln 154
              ::::::::::::::::|||||:::   |||:::|||      ||||||||   ||||||
Db        421 CTTCGGCTCATCGAACGGCTGGCGAAAGCGGCCGGTCCGGAAGGGATGGTCGCCGGTCAG 480

Qy        155 MetLeuAspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHis 174
              ||||||:::   ||||||:::   :::   |||      ||||||   ||||||
Db        481 GCAGCCGATATGGAAGGAGAGGGGAAAACGCTGACGCTTTCGGAGCTCGAATACATTCAT 540

Qy        175 LysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsn 194
              :::   ||||||||   :::|||   ::::::|||   :::   |||   |||
Db        541 CGGCATAAAACCGGGAAAATGCTGCAATACAGCGTGCACGCCGGCGCCTTGATCGGCGGC 600

Qy        195 ValAspAspThrThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPhe 214
              |||            |||:::   ::::::   |||||||||:::   |||
Db        601 GCTGATGCCCGGCAAACGCGGGAGCTTGACGAATTCGCCGCCCATCTAGGCCTTGCCTTT 660

Qy        215 GlnIleLysAspAspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysVal 234
              ||||||:::||||||:::||||||      |||   |||   |||:::||||||   |||
Db        661 CAAATTCGCGATGATATTCTCGATATTGAAGGGGCAGAAGAAAAAATCGGCAAGCCGGTC 720

Qy        235 GlySerAspLeuGluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAla 254
              |||||||||      ||||||||:::||||||   :::||||||            ||||||
Db        721 GGCAGCGACCAAAGCAACAACAAAGCGACGTATCCAGCGTTGCTGTCGCTTGCCGGCGCG 780

Qy        255 GluAspLysLeuThrTyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAsp 273
              ::::::||||||   :::|||   :::|||||         |||         |||
Db        781 AAGGAAAAGTTGGCGTTCCATATCGAGGCGGCGCAGCGCCATTTACGGAACGCCGAC 837

RESULT 6
US-08-534-910B-2
; Sequence 2, Application US/08534910B
; Patent No. 5766911
;  GENERAL INFORMATION:
;    APPLICANT:  KOIKE, Ayumi
;    APPLICANT:  OBATA, Shusei
;    APPLICANT:  NISHINO, Tokuzo
;    APPLICANT:  OHNUMA, Shinichi
;    APPLICANT:  NAKAZAWA, Takeshi
;    APPLICANT:  OGURA, Kyozo
;    APPLICANT:  KOYAMA, Tanetoshi
;    TITLE OF INVENTION:  Mutated Farnesyldiphosphate Synthase Capable
;    TITLE OF INVENTION:  Of Synthesizing Geranylgeranyldiphosphate And Gene Coding Ther
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenyon & Kenyon
;      STREET:  1025 Connecticut Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  U.S.
;      ZIP:  20036-5405
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.25" Floppy Disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;      SOFTWARE:  IBM/Word Perfect 6.1 Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/534,910B
;      FILING DATE:  28-SEPT-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-25253
;      FILING DATE:  14-FEB-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Toffenetti, Judith L.
;      REGISTRATION NUMBER:  39,048
;      REFERENCE/DOCKET NUMBER:  77670/398
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)429-1776
;      TELEFAX:  (202)429-0796
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  894 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Bacillus stearothermophilus
US-08-534-910B-2

Alignment Scores:
Pred. No.:                2.12e-71        Length:          894
Score:                    653.00          Matches:         132
Percent Similarity:       67.38%          Conservative:    56
Best Local Similarity:    47.31%          Mismatches:      85
Query Match:              44.94%          Indels:          6
DB:                       1               Gaps:            2

US-09-925-637-64 (1-287) x US-08-534-910B-2 (1-894)

Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              |||      |||   :::   :::   ::::::|||            :::      |||
Db          1 ATGGCGCAGCTTTCAGTTGAACAGTTTCTCAACGAGCAAAAACAGGCGGTGGAAACAGCG 60

Qy         21 IleAsnLysSerValMet----------AspThrGlnLeuGluGluSerMetLeuTyr 36
              :::::::::   :::                :::|||::::::::::|||   |||
Db         61 CTCTCCCGTTATATAGAGCGCTTAGAAGGGCCGGCGAAGGTGAAAAAGGCGATGGCGTAC 120

Qy         37 SerLeuAsnAlaGlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSer 56
              ||||||   ||||||||||||||||||||||:::|||||||||   |||:::   :::
Db        121 TCATTGGAGGCCGGCGGCAAACGAATCCGTCCGTTGCTGCTTCTGTCCACCGTTCAGGCG 180

Qy         57 LeuAsnThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThr 76
              |||      :::      :::|||:::      |||   |||:::|||||||||||||||
Db        181 CTCGGCAAAGACCCGGCGGTCGGATTGCCCGTCGCCTGCGCGATTGAAATGATCCATACG 240

Qy         77 TyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLys 96
              ||||||||||||||||||||||||||:::|||||||||||||||   ||||||||||||
Db        241 TACTCTTTGATCCATGATGATTTGCCGAGCATGGACAACGATGATTTGCGGCGCGGCAAG 300

Qy         97 LeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeu 116
              ||||||||||||||:::|||||||     ||||||||||||||||||   ||||||
Db        301 CCGACGAACCATAAAGTGTTCGGCGAGGCGATGGCCATCTTGGCGGGGGACGGGTTGTTG 360

Qy        117 ThrLysAlaPheGluLeuIleSer------SerAspAspArgLeuThrAspGluValLys 134
              |||   |||||:::|||||:::         |||:::|||:::         |||:::
Db        361 ACGTACGCGTTTCAATTGATCACCGAAATCGACGATGAGCGCATCCCTCCTTCCGTCCGG 420
```

```
Qy        135 IleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGln 154
              ::::::::::::::::|||||||:::   |||:::|||      |||||||||   ||||||
Db        421 CTTCGGCTCATCGAACGGCTGGCGAAAGCGGCCGGTCCGGAAGGGATGGTCGCCGGTCAG 480

Qy        155 MetLeuAspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHis 174
                  ||||||:::   ||||||:::   :::   |||      ||||||   ||||||
Db        481 GCAGCCGATATGGAAGGAGAGGGGAAAACGCTGACGCTTTCGGAGCTCGAATACATTCAT 540

Qy        175 LysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsn 194
              :::   |||||||||   :::|||   :::::::|||   :::   |||   |||
Db        541 CGGCATAAAACCGGGAAAATGCTGCAATACAGCGTGCACGCCGGCGCCTTGATCGGCGGC 600

Qy        195 ValAspAspThrThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPhe 214
                  |||                 |||:::   ::::::   ||||||||||:::   |||
Db        601 GCTGATGCCCGGCAAACGCGGGAGCTTGACGAATTCGCCGCCCATCTAGGCCTTGCCTTT 660

Qy        215 GlnIleLysAspAspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysVal 234
              ||||||:::||||||:::|||||||      |||   |||   |||:::|||||||   |||
Db        661 CAAATTCGCGATGATATTCTCGATATTGAAGGGGCAGAAGAAAAAATCGGCAAGCCGGTC 720

          235 GlySerAspLeuGluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAla 254
              |||||||||      |||||||||:::||||||   :::||||||         ||||||
Db        721 GGCAGCGACCAAAGCAACAACAAAGCGACGTATCCAGCGTTGCTGTCGCTTGCCGGCGCG 780

Qy        255 GluAspLysLeuThrTyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAsp 273
              ::::::||||||   :::|||   :::|||||||         |||      |||
Db        781 AAGGAAAAGTTGGCGTTCCATATCGAGGCGGCGCAGCGCCATTTACGGAACGCCGAC 837

RESULT 7
US-08-534-910B-3
; Sequence 3, Application US/08534910B
; Patent No. 5766911
;  GENERAL INFORMATION:
;    APPLICANT:  KOIKE, Ayumi
;    APPLICANT:  OBATA, Shusei
;    APPLICANT:  NISHINO, Tokuzo
;    APPLICANT:  OHNUMA, Shinichi
;    APPLICANT:  NAKAZAWA, Takeshi
;    APPLICANT:  OGURA, Kyozo
;    APPLICANT:  KOYAMA, Tanetoshi
;    TITLE OF INVENTION:  Mutated Farnesyldiphosphate Synthase Capable
;    TITLE OF INVENTION:  Of Synthesizing Geranylgeranyldiphosphate And Gene Coding Ther
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenyon & Kenyon
;      STREET:  1025 Connecticut Avenue, N.W., Suite 600
;      CITY:  Washington
       STATE:  DC
       COUNTRY:  U.S.
       ZIP:  20036-5405
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.25" Floppy Disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;      SOFTWARE:  IBM/Word Perfect 6.1 Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/534,910B
;      FILING DATE:  28-SEPT-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-25253
;      FILING DATE:  14-FEB-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Toffenetti, Judith L.
;      REGISTRATION NUMBER:  39,048
;      REFERENCE/DOCKET NUMBER:  77670/398
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)429-1776
;      TELEFAX:  (202)429-0796
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  894 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Bacillus stearothermophilus
US-08-534-910B-3

Alignment Scores:
Pred. No.:              4.97e-71        Length:         894
Score:                  650.00          Matches:        132
Percent Similarity:     67.03%          Conservative:   55
Best Local Similarity:  47.31%          Mismatches:     86
Query Match:            44.74%          Indels:         6
DB:                     1               Gaps:           2

US-09-925-637-64 (1-287) x US-08-534-910B-3 (1-894)

Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              |||      |||   :::   :::   :::::|||                :::    |||
Db          1 ATGGCGCAGCTTTCAGTTGAACAGTTTCTCAACGAGCAAAAACAGGCGGTGGAAACAGCG 60

Qy         21 IleAsnLysSerValMet-----------AspThrGlnLeuGluGluSerMetLeuTyr 36
              :::::::::   :::                      :::|||:::::::::|||   |||
Db         61 CTCTCCCGTTATATAGAGCGCTTAGAAGGGCCGGCGAAGCTGAAAAAGGCGATGGCGTAC 120

Qy         37 SerLeuAsnAlaGlyGlyLysArgIleArgProValLeuLeuLeuThrLeuAspSer 56
              ||||||   ||||||||||||||||||||||||:::|||||||||   |||:::   :::
Db        121 TCATTGGAGGCCGGCGGCAAACGAATCCGTCCGTTGCTGCTTCTGTCCACCGTTCGGGCG 180

Qy         57 LeuAsnThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThr 76
              |||      :::      :::|||:::      |||   |||:::|||||||||||||||
Db        181 CTCGGAAAAGACCCGGCGGTCGGATTGCCCGTCGCCTGCGCGATTGAAATGATCCATACG 240

Qy         77 TyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLys 96
              ||||||||||||||||||||||||||||:::|||||||||||||   ||||||||||
Db        241 TACTCTTTGATCCATGATGATTTGCCGAGCATGGACAACGATGATTTGCGGCGCGGCAAG 300

Qy         97 LeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeu 116
                 ||||||||||||||:::||||||      |||||||||||||||||   ||||||
Db        301 CCGACGAACCATAAAGTGTTCGGCGAGGCGATGGCCATCTTGGCGGGGGACGGGTTGTTG 360

Qy        117 ThrLysAlaPheGluLeuIleSer------SerAspAspArgLeuThrAspGluValLys 134
              |||   ||||||:::||||||:::         |||:::|||:::         |||:::
Db        361 ACGTACGCGTTTCAATTGATCACCGAAATCGACGATGAGCGCATCCCTCCTTCCGTCCGG 420

Qy        135 IleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGln 154
              ::::::::::::::::|||||||:::   |||:::|||      |||||||   ||||||
Db        421 CTTCGGCTCATCGAACGGCTGGCGAAAGCGGCCGGTCCGGAAGGGATGGCCGCCGGTCAG 480

Qy        155 MetLeuAspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHis 174
                  ||||||:::   ||||||:::   :::   |||      ||||||   ||||||
Db        481 GCAGCCGATATGGAAGGAGAGGGGAAAACGCTGACGCTTTCGGAGCTCGAATACATTCAT 540

Qy        175 LysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsn 194
              :::   |||||||||   :::|||   :::::::|||   :::   |||   |||
Db        541 CGGTATAAAACCGGGAAAATGCTGCAATACAGCGTGCACGCCGGCGCCTTGATCGGCGGC 600

Qy        195 ValAspAspThrThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPhe 214
                  |||                 |||:::   ::::::   ||||||||||:::   |||
Db        601 GCTGATGCCCGGCAAACGCGGGAGCTTGACGAATTCGCCGCCCATCTAGGCCTTGCCTTT 660

Qy        215 GlnIleLysAspAspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysVal 234
              ||||||:::||||||:::|||||||      |||   |||   |||:::|||||||   |||
Db        661 CAAATTCGCGATGATATTCTCGATATTGAAGGGGCAGAAGAAAAAATCGGCAAGCCGGTC 720

Qy        235 GlySerAspLeuGluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAla 254
              |||||||||      |||||||||:::||||||   :::||||||         ||||||
Db        721 GGCAGCGACCAAAGCAACAACAAAGCGACGTATCCAGCGTTGCTGTCGCTTGCCGGCGCA 780

Qy        255 GluAspLysLeuThrTyrHisArgAspAlaAlaValAspGluLeuThrGlnIleAsp 273
              ::::::||||||   :::|||   :::|||||||         |||      |||
Db        781 AAGGAAAAGTTGGCGTTCCATATCGAGGCGGCGCAGCGCCATTTACGGAACGCCGAC 837
```

```
RESULT 8
US-08-886-466-1
; Sequence 1, Application US/08886466C
; Patent No. 6040165
; GENERAL INFORMATION:
;  APPLICANT: Narita, Keishi
;  APPLICANT: Ishida, Chika
;  APPLICANT: Takeuchi, Yoshie
;  APPLICANT: Ohto, Chikara
;  APPLICANT: Ohnuma, Shinichi
;  APPLICANT: Nishino, Tokuzo
;  TITLE OF INVENTION: MUTANT PRENYL DIPHOSPHATE SYNTHASE
;  FILE REFERENCE: 77670/494
;  CURRENT APPLICATION NUMBER: US/08/886,466C
;  CURRENT FILING DATE: 1997-07-10
;  EARLIER APPLICATION NUMBER: JP 8-191635
;  EARLIER FILING DATE: 1996-07-03
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
    LENGTH: 894
    TYPE: DNA
;   ORGANISM: Bacillus stearothermophilus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(891)
US-08-886-466-1

Alignment Scores:
Pred. No.:              4.97e-71      Length:          894
Score:                  650.00        Matches:         134
Percent Similarity:     67.78%        Conservative:    49
Best Local Similarity:  49.63%        Mismatches:      85
Query Match:            44.74%        Indels:          2
DB:                     3             Gaps:            1

US-09-925-637-64 (1-287) x US-08-886-466-1 (1-894)

Qy          6 MetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerVal 25
              :::|||:::           |||         ||||||      |||   :::
Db         28 CTCAACGAGCAAAAACAGGCGGTGGAAACAGCGCTCTCCCGTTATATAGAGCGCTTAGAA 87

Qy         26 MetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIle 45
                 :::|||:::::::::|||   ||||||||   ||||||||||||||||||
Db         88 GGGCCGGCGAAGCTGAAAAAGGCGATGGCGTACTCATTGGAGGCCGGCGGCAAACGAATC 147

Qy         46 ArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMet 65
              ||||||:::|||||||||   |||:::   :::|||      :::      :::|||:::
Db        148 CGTCCGTTGCTGCTTCTGTCCACCGTTCGGGCGCTCGGCAAAGACCCGGCGGTCGGATTG 207

Qy         66 LysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
                ||| |||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 CCCGTCGCCTGCGCGATTGAAATGATCCATACGTACTCTTTGATCCATGATGATTTGCCG 267

Qy         86 AlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGlu 105
              :::|||||||||||||||   ||||||||||||   ||||||||||||||:::||||||
Db        268 AGCATGGACAACGATGATTTGCGGCGCGGCAAGCCGACGAACCATAAAGTGTTCGGCGAG 327

Qy        106 TrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSer--- 124
              ||||||||||||||||||   ||||||||   ||||||:::||||||:::
Db        328 GCGATGGCCATCTTGGCGGGGGACGGGTTGTTGACGTACGCGTTTCAATTGATCACCGAA 387

Qy        125 ---SerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIle 143
                 |||:::|||:::         |||||||||||||||||||||||||||:::
Db        388 ATCGACGATGAGCGCATCCCTCCTTCCGTCCGGCTTCGGCTCATCGAACGGCTGGCGAAA 447

Qy        144 AlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGln 163
              |||:::|||     |||||||||   ||||||     |||||||::   ||||||:::
Db        448 GCGGCCGGTCCGGAAGGGATGGTCGCCGGTCAGGCAGCCGATATGGAAGGAGAGGGGAAA 507

Qy        164 ProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThr 183
                 :::   |||      ||||||   |||||||:::   |||||||||   :::|||
Db        508 ACGCTGACGCTTTCGGAGCTCGAATACATTCATCGGCATAAAACCGGGAAAATGCTGCAA 567

Qy        184 PheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeu 203
              ::::::|||   :::   |||   |||         |||                  |||
Db        568 TACAGCGTGCACGCCGGCGCCTTGATCGGCGGCGCTGATGCCCGGCAAACGCGGGAGCTT 627

Qy        204 GluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCys 223
              :::   ::::::   |||||||||:::   |||||||||:::||||||:::||||||
Db        628 GACGAATTCGCCGCCCATCTAGGCCTTGCCTTTCAAATTCGCGATGATATTCTCGATATT 687

Qy        224 TyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSer 243
                 |||   |||   |||:::||||||   ||||||||||||      |||||||||:::
Db        688 GAAGGGGCAGAAGAAAAAATCGGCAAGCCGGTCGGCAGCGACCAAAGCAACAACAAAGCG 747

Qy        244 ThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAsp 263
              ||||||   :::||||||         ||||||:::::::||||||   :::|||  :::
Db        748 ACGTATCCAGCGTTGCTGTCGCTTGCCGGCGCGAAGGAAAAGTTGGCGTTCCATATCGAG 807

Qy        264 AlaAlaValAspGluLeuThrGlnIleAsp 273
              ||||||         |||         |||
Db        808 GCGGCGCAGCGCCATTTACGGAACGCCGAC 837

RESULT 9
US-09-475-304-1
; Sequence 1, Application US/09475304
; Patent No. 6225096
; GENERAL INFORMATION:
;  APPLICANT: Narita, Keishi
;  APPLICANT: Ishida, Chika
;  APPLICANT: Takeuchi, Yoshie
;  APPLICANT: Ohto, Chikara
;  APPLICANT: Ohnuma, Shinichi
;  APPLICANT: Nishino, Tokuzo
;  TITLE OF INVENTION: MUTANT PRENYL DIPHOSPHATE SYNTHASE
;  FILE REFERENCE: 77670/494
;  CURRENT APPLICATION NUMBER: US/09/475,304
;  CURRENT FILING DATE: 1999-12-30
;  EARLIER APPLICATION NUMBER: JP 8-191635
;  EARLIER FILING DATE: 1996-07-03
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 894
;   TYPE: DNA
;   ORGANISM: Bacillus stearothermophilus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(891)
US-09-475-304-1

Alignment Scores:
Pred. No.:              4.97e-71      Length:          894
Score:                  650.00        Matches:         134
Percent Similarity:     67.78%        Conservative:    49
Best Local Similarity:  49.63%        Mismatches:      85
Query Match:            44.74%        Indels:          2
DB:                     4             Gaps:            1

US-09-925-637-64 (1-287) x US-09-475-304-1 (1-894)

Qy          6 MetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerVal 25
              :::|||:::           |||         ||||||      |||   :::
Db         28 CTCAACGAGCAAAAACAGGCGGTGGAAACAGCGCTCTCCCGTTATATAGAGCGCTTAGAA 87

Qy         26 MetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIle 45
                 :::|||:::::::::|||   ||||||||   ||||||||||||||||||
Db         88 GGGCCGGCGAAGCTGAAAAAGGCGATGGCGTACTCATTGGAGGCCGGCGGCAAACGAATC 147

Qy         46 ArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMet 65
              ||||||:::|||||||||   |||:::   :::|||      :::      :::|||:::
Db        148 CGTCCGTTGCTGCTTCTGTCCACCGTTCGGGCGCTCGGCAAAGACCCGGCGGTCGGATTG 207

Qy         66 LysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
                ||| |||:::||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        208 CCCGTCGCCTGCGCGATTGAAATGATCCATACGTACTCTTTGATCCATGATGATTTGCCG 267

Qy         86 AlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGlu 105
              :::|||||||||||||||   ||||||||||   ||||||||||||||:::|||||||
Db        268 AGCATGGACAACGATGATTTGCGGCGCGGCAAGCCGACGAACCATAAAGTGTTCGGCGAG 327

Qy        106 TrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSer--- 124
              ||||||||||||||||||   |||||||||   ||||||:::|||||||:::
Db        328 GCGATGGCCATCTTGGCGGGGGACGGGTTGTTGACGTACGCGTTTCAATTGATCACCGAA 387

Qy        125 ---SerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIle 143
                 |||:::|||:::         |||:::::::::::::::::::::|||||||:::
Db        388 ATCGACGATGAGCGCATCCCTCCTTCCGTCCGGCTTCGGCTCATCGAACGGCTGGCGAAA 447

Qy        144 AlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGln 163
              |||:::|||      |||||||||   ||||||      ||||||:::   |||||||:::
Db        448 GCGGCCGGTCCGGAAGGGATGGTCGCCGGTCAGGCAGCCGATATGGAAGGAGAGGGGAAA 507

Qy        164 ProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThr 183
                 :::   |||      ||||||   |||||||:::   |||||||||   :::|||
             508 ACGCTGACGCTTTCGGAGCTCGAATACATTCATCGGCATAAAACCGGGAAAATGCTGCAA 567

Qy        184 PheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeu 203
              ::::::|||   :::   |||   |||         |||                  |||
Db        568 TACAGCGTGCACGCCGGCGCCTTGATCGGCGGCGCTGATGCCCGGCAAACGCGGGAGCTT 627

Qy        204 GluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCys 223
              :::   ::::::   ||||||||:::   |||||||||:::|||||:::||||||
Db        628 GACGAATTCGCCGCCCATCTAGGCCTTGCCTTTCAAATTCGCGATGATATTCTCGATATT 687

Qy        224 TyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSer 243
                 |||   |||   |||:::||||||   ||||||||||      ||||||||::
Db        688 GAAGGGGCAGAAGAAAAAATCGGCAAGCCGGTCGGCAGCGACCAAAGCAACAACAAAGCG 747

Qy        244 ThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAsp 263
              ||||||   :::||||||         ||||||:::::||||||   :::|||   :::
Db        748 ACGTATCCAGCGTTGCTGTCGCTTGCCGGCGCGAAGGAAAAGTTGGCGTTCCATATCGAG 807

Qy        264 AlaAlaValAspGluLeuThrGlnIleAsp 273
              ||||||         |||         |||
Db        808 GCGGCGCAGCGCCATTTACGGAACGCCGAC 837

RESULT 10
US-09-101-126-4
; Sequence 4, Application US/09101126
; Patent No. 6316216
; GENERAL INFORMATION:
   APPLICANT: OHTO, CHIKARA
   APPLICANT: NAKANE, HIROYUKI
   APPLICANT: NISHINO, TOKUZO
;  APPLICANT: OHNUMA, SHINICHI
;  APPLICANT: HIROOKA, KAZUTAKE
;  TITLE OF INVENTION: MUTATED PRENYL DIPHOSPHATE SYNTHASES
;  FILE REFERENCE: 77670/566
;  CURRENT APPLICATION NUMBER: US/09/101,126
;  CURRENT FILING DATE: 1999-04-27
;  EARLIER APPLICATION NUMBER: PCT/JP97/03921
;  EARLIER FILING DATE: 1997-10-29
;  EARLIER APPLICATION NUMBER: JP 8-307506
;  EARLIER FILING DATE: 1996-11-05
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 894
;   TYPE: DNA
;   ORGANISM: Bacillus stearothermophilus
;   FEATURE:
;   OTHER INFORMATION: 256-276 is an Asp-rich coding domain
US-09-101-126-4

Alignment Scores:
Pred. No.:              4.97e-71        Length:         894
Score:                  650.00          Matches:        134
Percent Similarity:     67.78%          Conservative:   49
Best Local Similarity:  49.63%          Mismatches:     85
Query Match:            44.74%          Indels:         2
DB:                     4               Gaps:           1

US-09-925-637-64 (1-287) x US-09-101-126-4 (1-894)

Qy          6 MetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerVal 25
              :::|||:::            |||         ||||||      |||   :::
Db         28 CTCAACGAGCAAAAACAGGCGGTGGAAACAGCGCTCTCCCGTTATATAGAGCGCTTAGAA 87

Qy         26 MetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIle 45
                 :::|||:::::::::|||   ||||||||   ||||||||||||||||||
Db         88 GGGCCGGCGAAGCTGAAAAAGGCGATGGCGTACTCATTGGAGGCCGGCGGCAAACGAATC 147

Qy         46 ArgProValLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMet 65
              ||||||:::|||||||||   |||:::   :::|||      :::      :::|||:::
Db        148 CGTCCGTTGCTGCTTCTGTCCACCGTTCGGGCGCTCGGCAAAGACCCGGCGGTCGGATTG 207

Qy         66 LysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
                 |||   |||:::|||||||||||||||||||||||||||||||||||||||||||
Db        208 CCCGTCGCCTGCGCGATTGAAATGATCCATACGTACTCTTTGATCCATGATGATTTGCCG 267

Qy         86 AlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGlu 105
              :::|||||||||||||||   ||||||||||   ||||||||||||||:::|||||||
Db        268 AGCATGGACAACGATGATTTGCGGCGCGGCAAGCCGACGAACCATAAAGTGTTCGGCGAG 327

Qy        106 TrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSer--- 124
              ||||||||||||||||||   |||||||||   ||||||:::|||||||:::
Db        328 GCGATGGCCATCTTGGCGGGGGACGGGTTGTTGACGTACGCGTTTCAATTGATCACCGAA 387

Qy        125 ---SerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIle 143
                 |||:::|||:::         |||:::::::::::::::::::::|||||||:::
Db        388 ATCGACGATGAGCGCATCCCTCCTTCCGTCCGGCTTCGGCTCATCGAACGGCTGGCGAAA 447

Qy        144 AlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGln 163
              |||:::|||      |||||||||   ||||||      ||||||:::   |||||||:::
Db        448 GCGGCCGGTCCGGAAGGGATGGTCGCCGGTCAGGCAGCCGATATGGAAGGAGAGGGGAAA 507

Qy        164 ProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThr 183
                 :::   |||      ||||||   |||||||:::   |||||||||   :::|||
Db        508 ACGCTGACGCTTTCGGAGCTCGAATACATTCATCGGCATAAAACCGGGAAAATGCTGCAA 567

Qy        184 PheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeu 203
              ::::::|||   :::   |||   |||         |||                  |||
Db        568 TACAGCGTGCACGCCGGCGCCTTGATCGGCGGCGCTGATGCCCGGCAAACGCGGGAGCTT 627

Qy        204 GluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCys 223
              :::   ::::::   ||||||||:::   |||||||||:::|||||:::||||||
Db        628 GACGAATTCGCCGCCCATCTAGGCCTTGCCTTTCAAATTCGCGATGATATTCTCGATATT 687

Qy        224 TyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSer 243
                 |||   |||   |||:::||||||   ||||||||||      ||||||||::
Db        688 GAAGGGGCAGAAGAAAAAATCGGCAAGCCGGTCGGCAGCGACCAAAGCAACAACAAAGCG 747

Qy        244 ThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAsp 263
              ||||||   :::||||||         ||||||:::::||||||   :::|||   :::
Db        748 ACGTATCCAGCGTTGCTGTCGCTTGCCGGCGCGAAGGAAAAGTTGGCGTTCCATATCGAG 807

Qy        264 AlaAlaValAspGluLeuThrGlnIleAsp 273
              ||||||         |||         |||
Db        808 GCGGCGCAGCGCCATTTACGGAACGCCGAC 837

RESULT 11
US-09-367-528A-4
; Sequence 4, Application US/09367528A
; Patent No. 6395525
; GENERAL INFORMATION:
;  APPLICANT:  TOYOTA JIDOSHA KABUSHIKI KAISHA
;  TITLE OF INVENTION:  Geranyl Diphosphate Synthetase Gene
;  FILE REFERENCE:  PH-586
;  CURRENT APPLICATION NUMBER: US/09/367,528A
;  CURRENT FILING DATE:  1999-08-16
```

```
;  PRIOR APPLICATION NUMBER:  JP97/346686
;  PRIOR FILING DATE:  1997-12-16
;  NUMBER OF SEQ ID NOS:  6
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 894
;   TYPE: DNA
;   ORGANISM: Bacillus stearothermophilus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(894)
US-09-367-528A-4

Alignment Scores:
Pred. No.:              4.97e-71      Length:          894
Score:                  650.00        Matches:         134
Percent Similarity:     67.78%        Conservative:    49
Best Local Similarity:  49.63%        Mismatches:      85
Query Match:            44.74%        Indels:          2
DB:                     4             Gaps:            1

-09-925-637-64 (1-287) x US-09-367-528A-4 (1-894)

Qy          6 MetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerVal 25
              :::|||:::           |||        ||||||     |||   :::
Db         28 CTCAACGAGCAAAAACAGGCGGTGGAAACAGCGCTCTCCCGTTATATAGAGCGCTTAGAA 87

Qy         26 MetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIle 45
                 :::|||:::::::::|||   ||||||||.|||||||||||||||||
Db         88 GGGCCGGCGAAGCTGAAAAAGGCGATGGCGTACTCATTGGAGGCCGGCGGCAAACGAATC 147

Qy         46 ArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMet 65
              ||||||:::|||||||||   |||:::   :::|||      :::      :::|||:::
Db        148 CGTCCGTTGCTGCTTCTGTCCACCGTTCGGGCGCTCGGCAAAGACCCGGCGGTCGGATTG 207

Qy         66 LysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
                 |||   |||:::||||||||||||||||||||||||||||||||||||||||||||
Db        208 CCCGTCGCCTGCGCGATTGAAATGATCCATACGTACTCTTTGATCCATGATGATTTGCCG 267

Qy         86 AlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGlu 105
              :::|||||||||||||||   ||||||||||||   ||||||||||||||:::||||||
Db        268 AGCATGGACAACGATGATTTGCGGCGCGGCAAGCCGACGAACCATAAAGTGTTCGGCGAG 327

Qy        106 TrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSer--- 124
                 ||||||||||||||||   ||||||||   ||||||:::||||||:::
Db        328 GCGATGGCCATCTTGGCGGGGGACGGGTTGTTGACGTACGCGTTTCAATTGATCACCGAA 387

Qy        125 ---SerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIle 143
                 |||:::|||:::         |||:::::::::::::::::|||||||:::
Db        388 ATCGACGATGAGCGCATCCCTCCTTCCGTCCGGCTTCGGCTCATCGAACGGCTGGCGAAA 447

Qy        144 AlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGln 163
              |||:::|||      |||||||||   ||||||      ||||||::   ||||||:::
Db        448 GCGGCCGGTCCGGAAGGGATGGTCGCCGGTCAGGCAGCCGATATGGAAGGAGAGGGGAAA 507

Qy        164 ProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThr 183
                 :::   |||      ||||||   ||||||:::   |||||||||   :::|||
Db        508 ACGCTGACGCTTTCGGAGCTCGAATACATTCATCGGCATAAAACCGGGAAAATGCTGCAA 567

Qy        184 PheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeu 203
              :::::|||   :::   |||   |||         |||                  |||
Db        568 TACAGCGTGCACGCCGGCGCCTTGATCGGCGGCGCTGATGCCCGGCAAACGCGGGAGCTT 627

Qy        204 GluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCys 223
              :::   ::::::   |||||||||:::   |||||||||:::||||||:::||||||
Db        628 GACGAATTCGCCGCCCATCTAGGCCTTGCCTTTCAAATTCGCGATGATATTCTCGATATT 687

Qy        224 TyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSer 243
                 |||   |||   |||:::|||||||   ||||||||||||      |||||||||:::
Db        688 GAAGGGGCAGAAGAAAAAATCGGCAAGCCGGTCGGCAGCGACCAAAGCAACAACAAAGCG 747

Qy        244 ThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAsp 263
              ||||||   :::||||||         ||||||::::::||||||   :::|||   :::
```

```
Db        748 ACGTATCCAGCGTTGCTGTCGCTTGCCGGCGCGAAGGAAAAGTTGGCGTTCCATATCGAG 807

Qy        264 AlaAlaValAspGluLeuThrGlnIleAsp 273
                 ||||||          |||          |||
Db        808 GCGGCGCAGCGCCATTTACGGAACGCCGAC 837

RESULT 12
US-08-534-910B-1
; Sequence 1, Application US/08534910B
; Patent No. 5766911
;  GENERAL INFORMATION:
;    APPLICANT:  KOIKE, Ayumi
;    APPLICANT:  OBATA, Shusei
;    APPLICANT:  NISHINO, Tokuzo
;    APPLICANT:  OHNUMA, Shinichi
;    APPLICANT:  NAKAZAWA, Takeshi
;    APPLICANT:  OGURA, Kyozo
;    APPLICANT:  KOYAMA, Tanetoshi
;    TITLE OF INVENTION:  Mutated Farnesyldiphosphate Synthase Capable
;    TITLE OF INVENTION:  Of Synthesizing Geranylgeranyldiphosphate And Gene Coding T
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenyon & Kenyon
;      STREET:  1025 Connecticut Avenue, N.W., Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  U.S.
;      ZIP:  20036-5405
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.25" Floppy Disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 6.2
;      SOFTWARE:  IBM/Word Perfect 6.1 Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/534,910B
;      FILING DATE:  28-SEPT-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-25253
;      FILING DATE:  14-FEB-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Toffenetti, Judith L.
;      REGISTRATION NUMBER:  39,048
;      REFERENCE/DOCKET NUMBER:  77670/398
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)429-1776
;      TELEFAX:  (202)429-0796
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  894 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Bacillus stearothermophilus
US-08-534-910B-1

Alignment Scores:
Pred. No.:              2.06e-70      Length:          894
Score:                  645.00        Matches:         130
Percent Similarity:     68.63%        Conservative:    56
Best Local Similarity:  47.97%        Mismatches:      79
Query Match:            44.39%        Indels:          6
DB:                     1             Gaps:            2

US-09-925-637-64 (1-287) x US-08-534-910B-1 (1-894)

Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              |||      |||   :::   :::   ::::::|||           :::      |||
Db          1 ATGGCGCAGCTTTCAGTTGAACAGTTTCTCAACGAGCAAAAACAGGCGGTGGAAACAGCG 60

Qy         21 IleAsnLysSerValMet----------AspThrGlnLeuGluGluSerMetLeuTyr 36
```

```
Db         61 CTCTCCCGTTATATAGAGCGCTTAGAAGGGCCGGCGAAGCTGAAAAAGGCGATGGCGTAC 120

Qy         37 SerLeuAsnAlaGlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSer 56

Db        121 TCATTGGAGGCCGGCGGCAAACGAATCCGTCCGTTGCTGCTTCTGTCCACCGTTCGGGCG 180

Qy         57 LeuAsnThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThr 76

Db        181 CTCGGAAAAGACCCGGCGGTCGGATTGCCCGTCGCCTGCGCGATTGAAATGATCCATACG 240

Qy         77 TyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLys 96

Db        241 CACTCTTTGATCCATGATGATTTGCCGAGCATGGACAACGATGATTTGCGGCGCGGCAAG 300

Qy         97 LeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeu 116

Db        301 CCGACGAACCATAAAGTGTTCGGCGAGGCGATGGCCATCTTGGCGGGGGACGGGTTGTTG 360

Qy        117 ThrLysAlaPheGluLeuIleSer------SerAspAspArgLeuThrAspGluValLys 134

Db        361 ACGTACGCGTTTCAATTGATCACCGAAATCGACGATGAGCGCATCCCTCCTTCCGTCCGG 420

Qy        135 IleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGln 154

Db        421 CTTCGGCTCATCGAACGGCTGGCGAAAGCGGCCGGTCCGGAAGGGATGGTCGCCGGTCAG 480

Qy        155 MetLeuAspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHis 174

Db        481 GCAGCCGATATGGAAGGAGAGGGGAAAACGCTGACGCTTTCGGAGCTCGAATACATTCAT 540

Qy        175 LysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsn 194

Db        541 CGGCATAAAACCGGGAAAATGCTGCAATACAGCGTGCACGCCGGCGCCTTGATCGGCGGC 600

Qy        195 ValAspAspThrThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPhe 214

Db        601 GCTGATGCCCGGCAAACGCGGGAGCTTGACGAATTCGCCGCCCATCTAGGCCTTGCCTTT 660

Qy        215 GlnIleLysAspAspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysVal 234

Db        661 CAAATTCGCGATGATATTCTCGATATTGAAGGGGCAGAAGAAAAAATCGGCAAGCCGGTC 720

Qy        235 GlySerAspLeuGluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAla 254

Db        721 GGCAGCGACCAAAGCAACAACAAAGCGACGTATCCAGCGTTGCTGTCGCTTGCCGGCGCG 780

Qy        255 GluAspLysLeuThrTyrHisArgAspAlaAla 265

Db        781 AAGGAAAAGTTGGCGTTCCATATCGAGGCGGCG 813

RESULT 13
US-09-367-528A-2
; Sequence 2, Application US/09367528A
; Patent No. 6395525
; GENERAL INFORMATION:
;  APPLICANT:  TOYOTA JIDOSHA KABUSHIKI KAISHA
;  TITLE OF INVENTION:  Geranyl Diphosphate Synthetase Gene
;  FILE REFERENCE:  PH-586
;  CURRENT APPLICATION NUMBER: US/09/367,528A
;  CURRENT FILING DATE:  1999-08-16
;  PRIOR APPLICATION NUMBER:  JP97/346686
;  PRIOR FILING DATE:  1997-12-16
;  NUMBER OF SEQ ID NOS:  6
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 894
;   TYPE: DNA
;   ORGANISM: Bacillus stearothermophilus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(894)
US-09-367-528A-2

Alignment Scores:
Pred. No.:                2.74e-70      Length:          894
Score:                    644.00        Matches:         133
Percent Similarity:       67.41%        Conservative:    49
Best Local Similarity:    49.26%        Mismatches:      86
Query Match:              44.32%        Indels:          2
DB:                       4             Gaps:            1

US-09-925-637-64 (1-287) x US-09-367-528A-2 (1-894)

Qy          6 MetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerVal 25

Db         28 CTCAACGAGCAAAAACAGGCGGTGGAAACAGCGCTCTCCCGTTATATAGAGCGCTTAGAA 87

Qy         26 MetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIle 45

Db         88 GGGCCGGCGAAGCTGAAAAAGGCGATGGCGTACTCATTGGAGGCCGGCGGCAAACGAATC 147

Qy         46 ArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMet 65

Db        148 CGTCCGTTGCTGCTTCTGTCCACCGTTCGGGCGCTCGGCAAAGACCCGGCGGTCGGATTG 207

Qy         66 LysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85

Db        208 CCCGTCGCCTGCGCGATTGAAATGATCCATACGTACTTCTTGATCCATGATGATTTGCCG 267

Qy         86 AlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGlu 105

Db        268 AGCATGGACAACGATGATTTGCGGCGCGGCAAGCCGACGAACCATAAAGTGTTCGGCGAG 327

Qy        106 TrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSer--- 124

Db        328 GCGATGGCCATCTTGGCGGGGGACGGGTTGTTGACGTACGCGTTTCAATTGATCACCGAA 387

Qy        125 ---SerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIle 143

Db        388 ATCGACGATGAGCGCATCCCTCCTTCCGTCCGGCTTCGGCTCATCGAACGGCTGGCGAAA 447

Qy        144 AlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGln 163

Db        448 GCGGCCGGTCCGGAAGGGATGGTCGCCGGTCAGGCAGCCGATATGGAAGGAGAGGGGAAA 507

Qy        164 ProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThr 183

Db        508 ACGCTGACGCTTTCGGAGCTCGAATACATTCATCGGCATAAAACCGGGAAAATGCTGCAA 567

Qy        184 PheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeu 203

Db        568 TACAGCGTGCACGCCGGCGCCTTGATCGGCGGCGCTGATGCCCGGCAAACGCGGGAGCTT 627

Qy        204 GluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCys 223

Db        628 GACGAATTCGCCGCCCATCTAGGCCTTGCCTTTCAAATTCGCGATGATATTCTCGATATT 687

Qy        224 TyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSer 243

Db        688 GAAGGGGCAGAAGAAAAAATCGGCAAGCCGGTCGGCAGCGACCAAAGCAACAACAAAGCG 747

Qy        244 ThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAsp 263

Db        748 ACGTATCCAGCGTTGCTGTCGCTTGCCGGCGCGAAGGAAAAGTTGGCGTTCCATATCGAG 807

Qy        264 AlaAlaValAspGluLeuThrGlnIleAsp 273

Db        808 GCGGCGCAGCGCCATTTACGGAACGCCGAC 837

RESULT 14
US-09-275-742-1
; Sequence 1, Application US/09275742
; Patent No. 6130069
; GENERAL INFORMATION:
;  APPLICANT: Wilding, Edwina Imogen
```

```
;  APPLICANT: Gwynn, Michael
;  TITLE OF INVENTION: ispA
;  FILE REFERENCE: GM10205
;  CURRENT APPLICATION NUMBER: US/09/275,742
;  CURRENT FILING DATE: 1999-03-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 876
;   TYPE: DNA
;   ORGANISM: Streptococcus pneumoniae
US-09-275-742-1

Alignment Scores:
Pred. No.:                8.07e-59      Length:          876
Score:                    551.00        Matches:         122
Percent Similarity:       64.89%        Conservative:    48
Best Local Similarity:    46.56%        Mismatches:      88
Query Match:              37.92%        Indels:          4
DB:                       3             Gaps:            3

-09-925-637-64 (1-287) x US-09-275-742-1 (1-876)

Qy         28 ThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgPro 47
              :::   |||   |||||||:::|||||||||:::::|||||||||||||||||||||||
Db         73 TCTAGTTTACGGGAGTCTGTTCTCTATTCTATTCATGCTGGTGGCAAGCGTATTCGGCCT 132

Qy         48 ValLeuLeuLeuLeuThrLeuAspSerLeuAsn---ThrGluTyrGluLeuGlyMetLys 66
                 ||||||||      |||::::::|||      |||                 :::
Db        133 TTTCTCTTGTTAAAAGTTCTGGAAGCCTTGCAGGTTACCATCAAACCTGCTCACGCGCAG 192

Qy         67 SerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAla 86
              |||   ||||||||||||||||||||   ||||||||||||||||||||||||||||
Db        193 GTAGCTACTGCCTTGGAGATGATTCATACAGGGAGCTTGATTCACGATGACCTTCCTGCT 252

Qy         87 MetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrp 106
              ||||||:::|||||||||||||||||:::|||||||||||||||   :::|||||||
Db        253 ATGGATGATGACGATTATCGAAGAGGGCGGTTAACCAATCACAAGAAATTCGGTGAAGCT 312

Qy        107 ThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAsp 126
              ||||||||||||||||||||||||            :::   ||||||:::
Db        313 ATGGCCATTTTGGCTGGAGATGCCTTATTCTTAGACCCATATGCCTTGATTGCGCAGGCA 372

Qy        127 AspArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGly 146
              |||   |||      ::::::|||:::   ::::::      ||||||:::|||||||||
Db        373 GAT---TTGCCAAGTCAGATCAAGGTGGGCTTGATTGCCAACTTATCCCTTGCTTCAGGT 429

Qy        147 HisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIleAsp 166
              :::||||||||||   ||||||:::||||||||:::   |||   |||   :::
Db        430 AGTCTGGGTATGGTGGCAGGGCAAGTTTTGGATATGGAGGGCGAACACCAGCACTTGTCT 489

Qy        167 LeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaVal 186
              ||||||   |||:::   ||||||      ||||||||   ||||||   :::
Db        490 CTGGAAGAACTTCAGACTATTCATGCCAATAAGACTGGGAAGTTACTAGCCTATCCCTTC 549

Qy        187 MetSerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyr 206
              :::|||||   ||||||  :::          :::      |||::::::
Db        550 CAAGCGGCAGCTATTATAGCTGAATTGTCACCTGAAATGCAGGTGAAGCTGAAAACTGTG 609

Qy        207 SerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGlyAsp 226
                  :::|||:::   ||||||:::::|||||||:::||||||
Db        610 GGTGAATTGATTGGACTTGCTTTTCAAGTCAGAGATGATGTACTGGATGTGACGGCTAGT 669

Qy        227 GluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyrVal 246
                 ::::::||||||                ||||||:::      ||||||||||||
Db        670 TTTGAGGAAATCGGCAAGACACCTCAAAAGGATCTGCAGGCAGAAAAATCAACCTATCCT 729

Qy        247 SerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAlaVal 266
              :::||||||||||   :::   :::                        |||   |||
Db        730 GCCTTGTTGGGCTTGGAAGAGTCCATTGCCTTTTGTAACCAAACCCTGGATCAAGCTAAT 789

Qy        267 AspGluLeuThrGlnIleAspGluGln------PheAsnThrLysHisLeuLeuGluIle 284
              |||:::|||   :::|||   :::|||      |||   |||:::   ::::::   :::
```

```
Db        790 GATAAATTAGAAGAAATTGCCCAGCAGCTTCCCTTTGAAACAGAATCGATTGTAAGTGTA 849

Qy        285 ValAsp 286
                 |||:::
Db        850 GTAGAA 855

RESULT 15
US-08-961-527-76/c
; Sequence 76, Application US/08961527
; Patent No. 6420135
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Streptococcus pneumoniae Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  391
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,527
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB340P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  76:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10011 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-961-527-76

Alignment Scores:
Pred. No.:                5.51e-57      Length:          10011
Score:                    549.00        Matches:         121
Percent Similarity:       65.27%        Conservative:    50
Best Local Similarity:    46.18%        Mismatches:      87
Query Match:              37.78%        Indels:          4
DB:                       4             Gaps:            3

US-09-925-637-64 (1-287) x US-08-961-527-76 (1-10011)

Qy         28 ThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgPro 47
              :::   |||   |||||||:::|||||||||:::::|||||||||||||||||||||||
Db       7806 TCTAGTTTACGGGAGTCTGTTCTCTATTCTATTCATGCTGGTGGCAAGCGTATTCGGCCT 7747

Qy         48 ValLeuLeuLeuLeuThrLeuAspSerLeuAsn---ThrGluTyrGluLeuGlyMetLys 66
                 ||||||||      |||::::::|||      |||                 :::
Db       7746 TTTCTCTTGTTAGAAGTTCTGGAAGCCTTGCAGGTTACCATCAAACCTGCTCACGCGCAG 7687

Qy         67 SerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAla 86
              |||   ||||||||||||||||||||   ||||||||||||||||||||||||||||
Db       7686 GTAGCTACTGCCTTAGAGATGATTCATACAGGGAGCTTGATTCACGATGACCTTCCTGCT 7627

Qy         87 MetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrp 106
              ||||||:::|||||||||||||||||:::|||||||||||||||   :::|||||||
Db       7626 ATGGATGATGACGATTATCGAAGAGGGCGGTTAACCAATCACAAGAAATTCGGTGAAGCT 7567
```

```
Qy        107 ThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAsp 126
              ||||||||||||||||||||||||          ::::::   |||||||:::
Db       7566 ATGGCCATTTTGGCTGGAGATGCCTTATTCTTAGACTCATATGCCTTGATTGCGCAGGCA 7507

Qy        127 AspArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGly 146
              |||   |||      :::::::|||:::   ::::::      |||||||:::|||||||||
Db       7506 GAT---TTGCCAAGTCAGATTAAGGTGGACTTGATTGCCAACTTATCCCTTGCTTCAGGT 7450

Qy        147 HisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIleAsp 166
              :::||||||||||   ||||||:::||||||||||:::   |||   |||   :::
Db       7449 AGTCTGGGTATGGTAGCAGGGCAAGTTTTGGATATGGAGGGCGAACACCAGCACTTGTCT 7390

Qy        167 LeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaVal 186
              ||||||   |||:::   ||||||      |||||||||   ||||||   :::
Db       7389 TTGGAAGAACTTCAGACTATTCATGCCAATAAGACTGGGAAGTTACTAGCCTATCCCTTC 7330

Qy        187 MetSerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyr 206
              :::||||||   ||||||   :::              :::      |||::::::
Db       7329 CAAGCGGCAGCTATTATAGCTGAATTGTCACCTGAAATGCAGGTGAAGCTGAAAACTGTG 7270

Qy        207 SerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGlyAsp 226
                    :::|||:::   ||||||::::::||||||:::||||||
Db       7269 GGTGAATTGATTGGACTTGCTTTTCAAGTCAGAGATGATGTACTGGATGTGACAGCTAGT 7210

Qy        227 GluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyrVal 246
                   ::::::||||||            ||||||:::      ||||||||||||
Db       7209 TTTGAGGAAATCGGCAAGACACCTCAAAAGGATCTGCAGGCAGAAAAATCAACCTATCCT 7150

Qy        247 SerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAlaVal 266
              :::|||||||||   :::   :::                        |||   |||
Db       7149 GCCCTGTTGGGCTTGGAAGAGTCCATTGCCTTTTGTAACCAAACCCTGGATCAAGCTAAT 7090

Qy        267 AspGluLeuThrGlnIleAspGluGln------PheAsnThrLysHisLeuLeuGluIle 284
              :::::::|||   :::|||   :::|||      |||   |||:::   ::::::   :::
Db       7089 GAAAAATTAGAAGAAATTGCCCAGCAGCTTCCCTTTGAAACAGAATCGATTGTAAGTGTA 7030

Qy        285 ValAsp 286
              |||:::
Db       7029 GTAGAA 7024
```

Search completed: May 30, 2003, 22:32:48
Job time : 64 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 30, 2003, 20:55:11 ; Search time 187 Seconds
(without alignments)
2069.416 Million cell updates/sec

Title: US-09-925-637-64
Perfect score: 1453
Sequence: 1 MTNLPMNKLIDEVNNELSVA.....ELTQIDEQFNTKHLLEIVDL 287

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

earched: 845702 seqs, 674182571 residues

Total number of hits satisfying chosen parameters: 1691404

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09925637/runat_23052003_174819_23502/app_query.fasta_1.455
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09925637_@CGN_1_1_57_@runat_23052003_174819_23502
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	Query Match %	Length	DB	ID	Description
	1	1453	100.0	861	9	US-10-084-205-63	Sequence 63, Appl
	2	1453	100.0	861	10	US-09-925-637-63	Sequence 63, Appl
c	3	1453	100.0	1893	7	US-08-781-986A-155	Sequence 155, App
	4	1442	99.2	882	10	US-09-815-242-8485	Sequence 8485, Ap
	5	1415	97.4	864	10	US-09-815-242-4184	Sequence 4184, Ap
c	6	690	47.5	413	10	US-09-815-242-2822	Sequence 2822, Ap
	7	635	43.7	811	10	US-09-974-300-1015	Sequence 1015, Ap
	8	596	41.0	882	10	US-09-815-242-6533	Sequence 6533, Ap
	9	587.5	40.4	7528	10	US-09-070-927A-55	Sequence 55, Appl
c	10	570	39.2	337	10	US-09-815-242-3275	Sequence 3275, Ap
	11	549	37.8	876	10	US-09-815-242-9499	Sequence 9499, Ap
	12	547	37.6	876	10	US-09-815-242-9175	Sequence 9175, Ap
	13	516.5	35.5	891	9	US-09-941-947A-19	Sequence 19, Appl
	14	516.5	35.5	891	10	US-09-934-903-13	Sequence 13, Appl
	15	516.5	35.5	891	10	US-09-934-868-71	Sequence 71, Appl
	16	504.5	34.7	1470	9	US-10-108-915-25	Sequence 25, Appl
	17	497.5	34.2	888	10	US-09-815-242-7142	Sequence 7142, Ap
	18	494.5	34.0	1441	9	US-10-108-915-21	Sequence 21, Appl
	19	487.5	33.6	1268	9	US-10-108-915-17	Sequence 17, Appl
	20	485.5	33.4	900	10	US-09-815-242-5972	Sequence 5972, Ap
	21	474.5	32.7	888	10	US-09-815-242-7873	Sequence 7873, Ap
	22	473	32.6	900	10	US-09-815-242-9986	Sequence 9986, Ap
	23	465.5	32.0	1131	10	US-09-934-778-1	Sequence 1, Appli
	24	380.5	26.2	1062	9	US-10-108-915-15	Sequence 15, Appl
	25	376.5	25.9	912	9	US-09-941-947A-25	Sequence 25, Appl
c	26	375	25.8	640681	10	US-09-790-988-1	Sequence 1, Appli
	27	373.5	25.7	912	10	US-09-815-242-7285	Sequence 7285, Ap
	28	372.5	25.6	912	10	US-09-815-242-7450	Sequence 7450, Ap
	29	356	24.5	973	9	US-10-108-915-13	Sequence 13, Appl
c	30	352	24.2	536165	9	US-09-939-964-1	Sequence 1, Appli
	31	332	22.8	951	9	US-10-166-037-3	Sequence 3, Appli
	32	328	22.6	498	10	US-09-974-300-5453	Sequence 5453, Ap
	33	328	22.6	973	9	US-10-108-915-11	Sequence 11, Appl
	34	321	22.1	498	10	US-09-974-300-5438	Sequence 5438, Ap
	35	311.5	21.4	969	10	US-09-815-242-8505	Sequence 8505, Ap
	36	311.5	21.4	1161	9	US-10-108-915-37	Sequence 37, Appl
	37	311	21.4	972	10	US-09-815-242-6224	Sequence 6224, Ap
	38	310	21.3	972	10	US-09-815-242-9636	Sequence 9636, Ap
	39	309.5	21.3	888	10	US-09-547-267-2	Sequence 2, Appli
	40	309.5	21.3	8625	9	US-09-920-923-1	Sequence 1, Appli
	41	309.5	21.3	11233	9	US-09-920-923-27	Sequence 27, Appl
	42	308.5	21.2	936	10	US-09-815-242-4260	Sequence 4260, Ap
c	43	297.5	20.5	1253	9	US-09-920-923-35	Sequence 35, Appl
c	44	297	20.4	2000	10	US-09-887-576-115	Sequence 115, App
	45	294	20.2	990	10	US-09-815-242-7045	Sequence 7045, Ap

ALIGNMENTS

RESULT 1
US-10-084-205-63
; Sequence 63, Application US/10084205
; Publication No. US20030049648A1
; GENERAL INFORMATION:
; APPLICANT: Choi, Gil
; TITLE OF INVENTION: 37 Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB515P1
; CURRENT APPLICATION NUMBER: US/10/084,205
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/151,933
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn ver. 3.1
; SEQ ID NO 63
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-084-205-63

Alignment Scores:
Pred. No.: 2.1e-161
Score: 1453.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 100.0%
Length: 861
Matches: 287
Conservative: 0
Mismatches: 0
Indels: 0

DB: 9 Gaps: 0

US-09-925-637-64 (1-287) x US-10-084-205-63 (1-861)

```
Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy         61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy         81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300

Qy        101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600

Qy        201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660

Qy        221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Qy        241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy        261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy        281 LeuLeuGluIleValAspLeu 287
              |||||||||||||||||||||
Db        841 TTATTAGAAATCGTTGATTTA 861
```

RESULT 2
US-09-925-637-63
; Sequence 63, Application US/09925637
; Patent No. US20020103338A1
; GENERAL INFORMATION:
; APPLICANT: Choi
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
; FILE REFERENCE: PB560
; CURRENT APPLICATION NUMBER: US/09/925,637
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/151,933
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 08/781,986
; PRIOR FILING DATE: 1997-01-03
; PRIOR APPLICATION NUMBER: US 08/956,171
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/009,861
; PRIOR FILING DATE: 1996-01-06
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-637-63

Alignment Score:			
Pred. No.:	2.18e-161	Length:	861
Score:	1453.00	Matches:	287
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-925-637-64 (1-287) x US-09-925-637-63 (1-861)

```
Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy         61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy         81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300

Qy        101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600

Qy        201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
```

```
Qy        221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Qy        241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy        261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy        281 LeuLeuGluIleValAspLeu 287
              |||||||||||||||||||||
Db        841 TTATTAGAAATCGTTGATTTA 861

RESULT 3
US-08-781-986A-155/c
; Sequence 155, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
;   APPLICANT: Charles Kunsch
;   TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;   NUMBER OF SEQUENCES: 5255
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Human Genome Sciences, Inc.
;     STREET: 9410 Key West Avenue
;     CITY: Rockville
;     STATE: Maryland
;     COUNTRY: USA
;     ZIP: 20850
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;     COMPUTER: HP Vectra 486/33
;     OPERATING SYSTEM: MSDOS version 6.2
;     SOFTWARE: ASCII Text
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/781,986A
;     FILING DATE:
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:
;     FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Benson, Bob
;     REGISTRATION NUMBER: 30,446
;     REFERENCE/DOCKET NUMBER: PB248PP
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (301) 309-8504
;     TELEFAX: (301) 309-8512
;  INFORMATION FOR SEQ ID NO: 155:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1893 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
US-08-781-986A-155

Alignment Scores:
Pred. No.:              6.85e-161     Length:          1893
Score:                  1453.00       Matches:         287
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            100.00%       Indels:          0
DB:                     7             Gaps:            0

US-09-925-637-64 (1-287) x US-08-781-986A-155 (1-1893)

Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1314 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 1255

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1254 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 1195

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1194 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 1135

Qy         61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1134 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 1075

Qy         81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1074 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 1015

Qy        101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1014 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 955

Qy        121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 895

Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        894 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 835

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        834 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 775

Qy        181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        774 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 715

Qy        201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        714 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 655

Qy        221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 595

Qy        241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 535

Qy        261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 475

Qy        281 LeuLeuGluIleValAspLeu 287
              |||||||||||||||||||||
Db        474 TTATTAGAAATCGTTGATTTA 454

RESULT 4
US-09-815-242-8485
; Sequence 8485, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT: Ohlsen, Kari L.
;  APPLICANT: Zyskind, Judith W.
;  APPLICANT: Wall, Daniel
;  APPLICANT: Trawick, John D.
;  APPLICANT: Carr, Grant J.
;  APPLICANT: Yamamoto, Robert T.
;  APPLICANT: Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION: Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE: 2001-03-21
```

```
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8485
;   LENGTH: 882
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(882)
US-09-815-242-8485

Alignment Scores:
Pred. No.:              4.42e-160      Length:          882
Score:                  1442.00        Matches:         285
Percent Similarity:     99.30%         Conservative:    0
Best Local Similarity:  99.30%         Mismatches:      2
Query Match:            99.24%         Indels:          0
DB:                     10             Gaps:            0

US-09-925-637-64 (1-287) x US-09-815-242-8485 (1-882)

Qy          1 MetThrAsnLeuProMetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         21 IleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120

Qy         41 GlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

          61 TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIle 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy         81 HisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCAT 300

Qy        101 LysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPhe 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        121 GluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGlnArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        181 LeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLys 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTATTAACATTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAA 600

Qy        201 GluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTTCAGATTAAAGATGATTTA 660

Qy        221 LeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsn 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTAGACTGCTATGGTGATGAAGCGAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Qy        241 AsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyr 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATAAAAGTACATACGTGAGTTTATTAGGAAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy        261 HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHis 280
              ||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        781 CATAGAGACGCAGCAGTGGATGAACTAATGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy        281 LeuLeuGluIleValAspLeu 287
              |||||||||||||||||||||
Db        841 TTATTAGAAATCGTTGATTTA 861

RESULT 5
US-09-815-242-4184
; Sequence 4184, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4184
;   LENGTH: 864
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-09-815-242-4184

Alignment Scores:
Pred. No.:              6.38e-157      Length:          864
Score:                  1415.00        Matches:         280
Percent Similarity:     99.29%         Conservative:    0
Best Local Similarity:  99.29%         Mismatches:      2
Query Match:            97.38%         Indels:          0
DB:                     10             Gaps:            0

US-09-925-637-64 (1-287) x US-09-815-242-4184 (1-864)

Qy          6 MetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerVal 25
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTA 60

Qy         26 MetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIle 45
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATC 120

Qy         46 ArgProValLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMet 65
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATG 180

Qy         66 LysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCA 240

Qy         86 AlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGlu 105
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCATAAAGTATATGGTGAG 300

Qy        106 TrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSer 125
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGT 360

Qy        126 AspAspArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSer 145
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGT 420

Qy        146 GlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIle 165
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATT 480

Qy        166 AspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAla 185
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCACTATTAACATTTGCG 540

Qy        186 ValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSer 205
                 ||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db        541 GTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAAGAACATTTAGAAAGT 600

Qy        206 TyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGly 225
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATAGTTATCATTTAGGTATGATGTTTCAGATTAAAGATGATTTATTAGACTGCTATGGT 660

Qy        226 AspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyr 245
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GATGAAGCGAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTACATAC 720

Qy        246 ValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAla 265
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTGAGTTTATTAGGAAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGACGCAGCA 780

Qy        266 ValAspGluLeuThrGlnIleAspGluGlnPheAsnThrLysHisLeuLeuGluIleVal 285
                 ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GTGGATGAACTAATGCAAATTGATGAACAATTCAATACAAAACACTTATTAGAAATCGTT 840

Qy        286 AspLeu 287
                 ||||||
Db        841 GATTTA 846

RESULT 6
US-09-815-242-2822/c
; Sequence 2822, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2822
;   LENGTH: 413
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-09-815-242-2822

Alignment Scores:
Pred. No.:                3.34e-72        Length:          413
Score:                    690.00          Matches:         137
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              47.49%          Indels:          0
DB:                       10              Gaps:            0

US-09-925-637-64 (1-287) x US-09-815-242-2822 (1-413)

Qy         54 LeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMet 73
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 TTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATG 354

Qy         74 IleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArg 93
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 ATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGA 294

Qy         94 ArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAsp 113
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 CGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGAT 234

Qy        114 AlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAspArgLeuThrAspGluVal 133
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 GCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTA 174

Qy        134 LysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGly 153
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 AAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGT 114

Qy        154 GlnMetLeuAspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIle 173
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 CAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATA 54

Qy        174 HisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAla 190
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         53 CACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCA 3

RESULT 7
US-09-974-300-1015
; Sequence 1015, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
;  APPLICANT: Berka, Randy M.
;  APPLICANT:  Clausen, Ib Groth
;  TITLE OF INVENTION: Methods For Monitoring Multiple Gene
;  TITLE OF INVENTION:  Expression
```

```
;  FILE REFERENCE: 10085.500-US
;  CURRENT APPLICATION NUMBER: US/09/974,300
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: 09/680,598
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/279,526
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 8481
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1015
;   LENGTH: 811
;   TYPE: DNA
;   ORGANISM: Bacillus licheniformis
US-09-974-300-1015

Alignment Scores:
Pred. No.:              2.58e-65      Length:          811
Score:                  635.00        Matches:         125
Percent Similarity:     73.11%        Conservative:    49
Best Local Similarity:  52.52%        Mismatches:      62
Query Match:            43.70%        Indels:          2
DB:                     10            Gaps:            1

US-09-925-637-64 (1-287) x US-09-974-300-1015 (1-811)

Qy         30 LeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgProValLeu 49
              |||:::|||||||||||||||||||||||   |||||||||||||||:::|||||||:::|||
Db         97 CTCAAGGAATCCATGCTGTACTCATTGGAAGCCGGGGGGAAACGGCTGAGACCGATATTG 156

Qy         50 LeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSerAlaIle 69
              :::|||      |||   :::         :::   |||   |||:::
Db        157 GTTCTTGCGCTTCTTCATGCGTACGGAAAAGACGAAGAAGCCGGAATTCCAGTCGGATGC 216

Qy         70 AlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsn 89
              |||:::|||||||||||||||||||||||||||||||||||||||||   ||||||:::
Db        217 GCGGTTGAAATGATCCATACGTATTCATTAATCCATGACGACCTCCCTTGTATGGATGAC 276

Qy         90 AspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrpThrAlaIle 109
              ||||||   ||||||||||||   |||||||||||:::|||||||||   |||||||||
Db        277 GATGACCTCAGAAGAGGTAAGCCGACAAACCACAAAATTTACGGGGAGGCGACGGCCATT 336

Qy        110 LeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAsp------Asp 127
              ||||||||||||||||||||||:::::::|||:::::|||:::|||:::
Db        337 CTCGCGGGAGACGCCTTATTAACCGAAAGCTTCAAAATGATCACTTCCAATATGCCTTCT 396

Qy        128 ArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHis 147
              ::::::   |||   :::|||:::::::::      |||      |||:::|||
Db        397 GATGTATCCGCTGAAAAGCGCATCAGACTGGTGAACGAGCTGATTTCGGCAGCGGGGGCC 456

Qy        148 ValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIleAspLeu 167
              ||||||||||||||||||:::|||||||||:::::|||   :::   :::   |||
Db        457 GAAGGCATGGTCGGCGGTCAGATTTTAGATATGGAAGCGGAATCAAAGTCTGTATCTCTT 516

Qy        168 GluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMet 187
              :::   |||:::   ||||||:::   ||||||      |||||||:::|||:::|||:::
Db        517 GATGAACTGCAGCGAATCCACGAAGGAAAAACGGCTAAACTCCTCAGCTTCAGCGTCATT 576

Qy        188 SerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyrSer 207
              :::   |||   :::|||:::      :::      |||   |||      :::|||
Db        577 GCGGGAGCCATTCTTGCAGATGCATCTGAAAAAGAGATCGAAAAGCTGCGCGAATTCAGC 636

Qy        208 TyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGlyAspGlu 227
              :::|||:::|||:::   |||||||||:::||||||:::||||||      |||   |||
Db        637 CATCATATCGGCATCGGCTTTCAGATCAGAGATGACATTTTGGACCTTGAAGGCTCGGAG 696

Qy        228 AlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyrValSer 247
              |||:::|||||:::::|||||||||      |||   ||||||||||||   |||
Db        697 GATAAGATCGGCAAACGGATCGGATCCGACGCTTCAAACGGAAAGTCGACGTATCCGTCG 756

Qy        248 LeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAla 265
              ||||||     :::|||||||     ||||||      |||   :::   |||
Db        757 CTCCTTTCGCTTGAGGGCGCCAGCCAAAAACTTGATGAACATATTGAAAAAGCA 810
```

```
RESULT 8
US-09-815-242-6533
; Sequence 6533, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6533
;   LENGTH: 882
;   TYPE: DNA
;   ORGANISM: Enterococcus faecalis
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(882)
US-09-815-242-6533

Alignment Scores:
Pred. No.:              1.12e-60      Length:          882
Score:                  596.00        Matches:         134
Percent Similarity:     65.29%        Conservative:    56
Best Local Similarity:  46.05%        Mismatches:      91
Query Match:            41.02%        Indels:          10
DB:                     10            Gaps:            5

US-09-925-637-64 (1-287) x US-09-815-242-6533 (1-882)

Qy          1 MetThrAsnLeuProMetAsn------LysLeuIleAspGluValAsnAsnGluLeuSer 18
              |||||||||                   |||:::::::::|||      :::
Db          1 ATGACGAATTTTAGTCAACAGCATTTACCGTTGGTTGAAAAAGTCATGGTCGATTTT--- 57

Qy         19 ValAlaIleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeu 38
                 |||   :::         :::   :::|||:::|||:::|||||||||||:::
Db         58 ------ATCGCAGAATATACTGAAAATGAGCGTTTGAAAGAAGCAATGCTCTACTCAATA 111

Qy         39 AsnAlaGlyGlyLysArgIleArgProValLeuLeuLeuThrLeuAspSerLeuAsn 58
              :::|||||||||||||||:::||||||:::|||:::|||   |||:::   :::
Db        112 CACGCAGGTGGCAAACGATTACGCCCGCTATTAGTGCTAACAACAGTGGCCGCTTTTCAA 171

Qy         59 ThrGluTyrGluLeu---GlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyr 77
              |||   |||         :::   |||   :::|||||||||||||||||||||
Db        172 AAAGAGATGGAAACGCAAGACTATCAAGTGGCTGCCTCTTTAGAGATGATTCATACGTAT 231

Qy         78 SerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeu 97
              ||||||||||||||||||||||||||||||:::||||||   ||||||||||
Db        232 TCATTAATTCATGATGATTTACCAGCAATGGATGATGATGATTTACGTCGTGGCAAACCA 291
```

```
Qy        98 ThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThr 117
Db       292 ACCAATCACAAAGTGTTTGGTGAAGCGACTGCCATTTTAGCTGGGGACGGCTTATTAACA 351

Qy       118 LysAlaPheGluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysVal 137
Db       352 GGCGCATTTCAGTTACTTTCTTTGAGCCAATTAGGCTTAAGTGAA---AAAGTTTTACTG 408

Qy       138 LeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAsp 157
Db       409 ATGCAACAACTGGCAAAGGCAGCGGGGAATCAAGGCATGGTTGCCGGCCAAATGGGTGAT 468

Qy       158 MetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLys 177
Db       469 ATTGAAGGAGAAAAAGTCAGCTTGACTTTAGAAGAGTTAGCGGCTGTTCATGAAAAGAAA 528

Qy       178 ThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAsp 197
Db       529 ACCGGAGCGCTAATTGAATTCGCCTTAATTGCTGGGGGCGTATTAGCCAATCAAACAGAA 588

Qy       198 ThrThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLys 217
Db       589 GAAGTCATTGGTTTACTGACACAGTTTGCGCATCACTATGGCTTGGCTTTTCAAATTCGT 648

Qy       218 AspAspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAsp 237
Db       649 GATGACTTATTAGATGCCACAAGTACAGAAGCAGATTTAGGTAAAAAAGTAGGCCGAGAT 708

Qy       238 LeuGluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLys 257
Db       709 GAAGCGTTGAATAAAAGTACGTATCCAGCTCTTCTAGGGATTGCTGGCGCGAAAGATGCG 768

Qy       258 LeuThrTyr---------HisArgAspAlaAlaValAspGluLeuThrGlnIleAspGlu 274
Db       769 CTAACACATCAATTAGCAGAAGGCAGCGCTGTTTTAGAAAAAATCAAAGCAAATGTTCCC 828

Qy       275 GlnPheAsnThrLysHisLeuLeuGluIleVal 285
Db       829 AACTTTTCAGAAGAACATTTAGCAAACTTGTTA 861

RESULT 9
US-09-070-927A-55
; Sequence 55, Application US/09070927A
; Patent No. US20020120116A1
;   GENERAL INFORMATION:
;       APPLICANT: Charles A. Kunsch
;                  Patrick J. Dillon
;                  Steven Barash
;       TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
;       NUMBER OF SEQUENCES: 982
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Human Genome Sciences, Inc.
;            STREET: 9410 Key West Avenue
;            CITY: Rockville
;            STATE: Maryland
;            COUNTRY: USA
;            ZIP: 20850
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;            COMPUTER: HP Vectra 486/33
;            OPERATING SYSTEM: MSDOS version 6.2
;            SOFTWARE: ASCII Text
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/070,927A
;            FILING DATE: 04-May-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 60/046,655
;            FILING DATE: 1997-05-16
;            APPLICATION NUMBER: 60/044,031
;            FILING DATE: 1997-05-06
;            APPLICATION NUMBER: 60/066,009
;            FILING DATE: 1997-11-14
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Kenley K. Hoover
;            REGISTRATION NUMBER: 40,302
;            REFERENCE/DOCKET NUMBER: PB369
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (301) 309-8504
;            TELEFAX: (301) 309-8512
;   INFORMATION FOR SEQ ID NO: 55:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 7528 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 55:
US-09-070-927A-55

Alignment Scores:
Pred. No.:              2.52e-58      Length:          7528
Score:                  587.50        Matches:         130
Percent Similarity:     67.68%        Conservative:    48
Best Local Similarity:  49.43%        Mismatches:      78
Query Match:            40.43%        Indels:          7
DB:                     10            Gaps:            4

US-09-925-637-64 (1-287) x US-09-070-927A-55 (1-7528)

Qy         1 MetThrAsnLeuProMetAsn------LysLeuIleAspGluValAsnAsnGluLeuSer 18
Db      6642 ATGACGAATTTTAGTCAACAGCATTTACCGTTGGTTGAAAAAGTCATGGTCGATTTT--- 6698

Qy        19 ValAlaIleAsnLysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeu 38
Db      6699 ------ATCGCAGAATATACTGAAAATGAGCGTTTGAAAGAAGCAATGCTCTACTCAATA 6752

Qy        39 AsnAlaGlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsn 58
Db      6753 CACGCAGGTGGCAAACGATTACGCCCGCTATTAGTGTTAACAACAGTGGCCGCTTTTCAA 6812

Qy        59 ThrGluTyrGluLeu---GlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyr 77
Db      6813 AAAGAGATGGAAACACAAGACTATCAAGTGGCTGCCTCTTTAGAGATGATTCATACGTAT 6872

Qy        78 SerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeu 97
Db      6873 TCATTAATTCATGATGATTTACCAGCAATGGACGATGATGATTTACGTCGTGGCAAACCA 6932

Qy        98 ThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThr 117
Db      6933 ACCAATCATAAAGTGTTTGGTGAAGCGACTGCCATTTTAGCAGGGGACGGCTTATTAACA 6992

Qy       118 LysAlaPheGluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysVal 137
Db      6993 GGTGCATTTCAGTTGCTTTCTTTGAGTCAATTAGGCTTAAGTGAA---AAAGTTTTACTG 7049

Qy       138 LeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAsp 157
Db      7050 ATGCAACAACTGGCAAAGGCAGCGGGGAATCAAGGCATGGTTTCCGGCCAAATGGGTGAT 7109

Qy       158 MetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLys 177
Db      7110 ATTGAAGGAGAAAAAGTCAGCTTGACTTTAGAAGAGTTAGCGGCTGTTCATGAAAAGAAA 7169

Qy       178 ThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAsp 197
Db      7170 ACCGGAGCGCTAATTGAATTCGCCTTAATTGCTGGGGGCGTATTAGCCAATCAAACAGAA 7229

Qy       198 ThrThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLys 217
Db      7230 GAAGTCATTGGTTTACTGACACAGTTTGCGCATCACTATGGCTTGGCTTTTCAAATTCGT 7289

Qy       218 AspAspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAsp 237
Db      7290 GATGACTTATTAGATGCCACAAGTACAGAAGCAGATTTAGGTAAAAAAGTAGGCCGAGAT 7349
```

```
Qy        238 LeuGluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLys 257
              ||||||||||||||||  :::|||||||||      ||||||:::|||
Db       7350 GAAGCGTTGAATAAAAGTACGTATCCAGCTCTTCTAGGGATTGCTGGCGCGAAAGATGCG 7409

Qy        258 LeuThrTyr 260
              |||||:::
Db       7410 CTAACACAT 7418

RESULT 10
US-09-815-242-3275/c
; Sequence 3275, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
   APPLICANT:  Carr, Grant J.
   APPLICANT:  Yamamoto, Robert T.
   APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3275
;   LENGTH: 337
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-09-815-242-3275

Alignment Scores:
Pred. No.:                3.12e-58        Length:          337
Score:                    570.00          Matches:         112
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              39.23%          Indels:          0
DB:                       10              Gaps:            0

US-09-925-637-64 (1-287) x US-09-815-242-3275 (1-337)

Qy         76 ThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGly 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        336 ACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGA 277

Qy         96 LysLeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeu 115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 AAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTA 217

Qy        116 LeuThrLysAlaPheGluLeuIleSerSerAspAspArgLeuThrAspGluValLysIle 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 TTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATA 157

Qy        136 LysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMet 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        156 AAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATG 97

Qy        156 LeuAspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLys 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 TTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAA 37

Qy        176 ThrLysThrGlyAlaLeuLeuThrPheAlaValMet 187
              |||||||||||||||||||||||||||||||||||||
Db         36 ACAAAAACAGGAGCATTATTAACTTTTGCGGTTATG 1

RESULT 11
US-09-815-242-9499
; Sequence 9499, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9499
;   LENGTH: 876
;   TYPE: DNA
;   ORGANISM: Streptococcus pneumoniae
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(876)
US-09-815-242-9499

Alignment Scores:
Pred. No.:                3.68e-55        Length:          876
Score:                    549.00          Matches:         121
Percent Similarity:       65.27%          Conservative:    50
Best Local Similarity:    46.18%          Mismatches:      87
Query Match:              37.78%          Indels:          4
DB:                       10              Gaps:            3

US-09-925-637-64 (1-287) x US-09-815-242-9499 (1-876)

Qy         28 ThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgPro 47
              :::   |||   ||||||:::|||||||||::::::|||||||||||||||||||||||
Db         73 TCTAGTTTACGGGAGTCTGTTCTCTATTCTATTCATGCTGGTGGCAAGCGTATTCGGCCT 132

Qy         48 ValLeuLeuLeuLeuThrLeuAspSerLeuAsn---ThrGluTyrGluLeuGlyMetLys 66
              |||||||||      |||:::::|||      |||                    :::
Db        133 TTTCTCTTGTTAGAAGTTCTGGAAGCCTTGCAGGTTACCATCAAACCTGCTCACGCGCAG 192

Qy         67 SerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAla 86
              |||   ||||||||||||||||||||   ||||||||||||||||||||||||||||||
```

```
Db        193 GTAGCTACTGCCTTGGAGATGATTCATACAGGGAGCTTGATTCACGATGACCTTCCTGCT 252

Qy         87 MetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrp 106
              ||||||:::|||||||||||||||||||:::|||||||||||||||   :::||||||
Db        253 ATGGATGATGACGATTATCGAAGAGGGCGGTTAACCAATCACAAGAAATTCGGTGAAGCT 312

Qy        107 ThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAsp 126
              ||||||||||||||||||||||||          :::   ||||||:::
Db        313 ATGGCCATTTTGGCTGGAGATGCCTTATTCTTAGACCCATATGCCTTGATTGCGCAGGCA 372

Qy        127 AspArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGly 146
              |||   |||      :::::::|||:::   ::::::      ||||||:::||||||||
Db        373 GAT---TTGCCAAGTCAGATTAAGGTGGACTTGATTGCCAACTTATCCCTTGCTTCAGGT 429

Qy        147 HisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIleAsp 166
              :::||||||||||   ||||||:::||||||||:::   |||   |||   :::
Db        430 AGTCTGGGTATGGTGGCAGGGCAAGTTTTGGATATGGAGGGCGAACACCAGCACTTGTCT 489

Qy        167 LeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaVal 186
              ||||||   |||:::::||||||      ||||||||   ||||||   :::
Db        490 TTGGAAGAACTTCAGATTATTCATGCCAATAAGACTGGGAAGTTACTAGCCTATCCCTTC 549

Qy        187 MetSerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyr 206
              :::||||||   ||||||  ::::          :::      |||:::::
Db        550 CAAGCGGCAGCTATTATAGCTGAATTGTCACCTGAAATGCAGGTGAAGCTGAAAACTGTG 609

Qy        207 SerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGlyAsp 226
                  :::|||:::   ||||||:::::||||||:::||||||
Db        610 GGTGAATTGATTGGACTTGCTTTTCAAGTCAGAGATGATGTACTGGATGTGACAGCTAGT 669

Qy        227 GluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyrVal 246
                  ::::::||||||          ||||||:::      ||||||||||||
Db        670 TTTGAGGAAATCGGCAAGACACCTCAAAAGGATCTGCAGGCAGAAAAATCAACCTATCCT 729

Qy        247 SerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAlaVal 266
              :::||||||||||   :::   :::                        |||   |||
Db        730 GCCCTGTTGGGCTTGGAAGAGTCCATTGCCTTTTGTAACCAAACCCTGGATCAAGCTAAT 789

Qy        267 AspGluLeuThrGlnIleAspGluGln------PheAsnThrLysHisLeuLeuGluIle 284
              :::::|||   :::|||   :::|||      |||   |||:::   ::::::   :::
Db        790 GAAAAATTAGAAGAAATTGCCCAGCAGCTTCCCTTTGAAACAGAATCGATTGTAAGTGTA 849

Qy        285 ValAsp 286
              |||:::
Db        850 GTAGAA 855

RESULT 12
US-09-815-242-9175
; Sequence 9175, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9175
;   LENGTH: 876
;   TYPE: DNA
;   ORGANISM: Streptococcus pneumoniae
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(876)
US-09-815-242-9175

Alignment Scores:
Pred. No.:              6.32e-55        Length:         876
Score:                  547.00          Matches:        120
Percent Similarity:     64.89%          Conservative:   50
Best Local Similarity:  45.80%          Mismatches:     88
Query Match:            37.65%          Indels:         4
DB:                     10              Gaps:           3

US-09-925-637-64 (1-287) x US-09-815-242-9175 (1-876)

Qy         28 ThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgPro 47
              :::   |||   ||||||:::||||||||:::::||||||||||||||||||||||||
Db         73 TCTAGTTTACGGGAGTCTGTTCTCTATTCTATTCATGCTGGTGGCAAGCGTATTCGGCCT 132

Qy         48 ValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGlu---LeuGlyMetLys 66
              ||||||||   |||:::::|||          :::          :::
Db        133 TTTCTCTTGTTAGAAGTTCTGGAAGCCTTGCAGGTTGCCATCAAACCTGCTCACGCGCAG 192

Qy         67 SerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAla 86
              |||   ||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db        193 GTAGCTACTGCCTTGGAGATGATTCATACAGGGAGCTTGATTCACGATGACCTTCCTGCT 252

Qy         87 MetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrp 106
              ||||||:::|||||||||||||||||||:::|||||||||||||||   :::||||||
Db        253 ATGGATGATGACGATTATCGAAGAGGGCGGTTAACCAATCACAAGAAATTCGGTGAAGCT 312

Qy        107 ThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAsp 126
              ||||||||||||||||||||||||          :::   ||||||:::
Db        313 ATGGCCATTTTGGCTGGAGATGCCTTATTCTTAGACCCATATGCCTTGATTGCGCAGGCA 372

Qy        127 AspArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGly 146
              |||   |||      :::::::|||:::   ::::::      ||||||:::||||||||
Db        373 GAT---TTGCCAAGTCAGATCAAGGTGGACTTGATTGCCAACTTATCCCTTGCTTCAGGT 429

Qy        147 HisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIleAsp 166
              :::||||||||||   ||||||:::||||||||:::   |||   |||   :::
Db        430 AGTCTGGGTATGGTGGCAGGGCAAGTTTTGGATATGGAGGGCGAACACCAGCACTTGTCT 489

Qy        167 LeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaVal 186
              ||||||   |||:::   ||||||      ||||||||   ||||||   :::
Db        490 CTGGAAGAACTTCAGACTATTCATGCCAATAAGACTGGGAAGTTACTAGCCTATCCCTTC 549

Qy        187 MetSerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyr 206
              :::||||||   ||||||  :::            :::      |||:::::
Db        550 CAAGCGGCAGCTATTATAGCTGAATTGTCACCTGAAATGCAGGTGAAGCTGAAAACTGTG 609

Qy        207 SerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGlyAsp 226
                  :::|||:::   ||||||:::::||||||:::||||||
Db        610 GGTGAATTGATTGGACTTGCTTTTCAAGTCAGAGATGATGTACTGGATGTGACAGCTAGT 669

Qy        227 GluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyrVal 246
                  ::::::||||||          ||||||:::      ||||||||||||
Db        670 TTTGAGGAAATCGGTAAGACACCTCAAAAGGATCTGCAGGCAGAAAAATCAACCTATCCT 729

Qy        247 SerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAlaVal 266
              :::||||||||||   :::   :::                        |||   |||
Db        730 GCCTTGTTGGGCTTGGAAGAGTCCATTGCCTTTTGTAACCAAACCCTGGATCAAGCTAAT 789
```

```
Qy        267 AspGluLeuThrGlnIleAspGluGln------PheAsnThrLysHisLeuLeuGluIle 284
              :::::::|||   :::|||   :::|||      |||   |||:::   ::::::   :::
Db        790 GAAAAATTGGAAGAAATTGCCCAGCAGGTTCCCTTTGAAACAGAATCGATTGTAAGTGTA 849

Qy        285 ValAsp 286
              |||:::
Db        850 GTAGAA 855

RESULT 13
US-09-941-947A-19
; Sequence 19, Application US/09941947A
; Publication No. US20030003528A1
; GENERAL INFORMATION:
;  APPLICANT: Brzostowicz, Patricia C.
;  APPLICANT:  Cheng, Qiong
;  APPLICANT:  DiCosimo, Deana J.
;  APPLICANT:  Koffas, Mattheos
;  APPLICANT:  Miller, Edward S. Jr.
;  APPLICANT:  Odom, J. Martin
;  APPLICANT:  Picataggio, Steve
;  APPLICANT:  Rouviere, Pierre E.
;  TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
;  FILE REFERENCE: CL1903 US NA
;  CURRENT APPLICATION NUMBER: US/09/941,947A
;  CURRENT FILING DATE:  2001-09-01
;  PRIOR APPLICATION NUMBER: 60/229,907
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/229,858
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
;   LENGTH: 891
;   TYPE: DNA
;   ORGANISM: Methylomonas 16a
US-09-941-947A-19

Alignment Scores:
Pred. No.:              2.48e-51      Length:          891
Score:                  516.50        Matches:         113
Percent Similarity:     63.75%        Conservative:    47
Best Local Similarity:  45.02%        Mismatches:      88
Query Match:            35.55%        Indels:          3
DB:                     9             Gaps:            2

US-09-925-637-64 (1-287) x US-09-941-947A-19 (1-891)

Qy         23 LysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGly 42
              ::::::::::::      |||   ::::::|||   |||||||::      ||||||
Db         76 GAAAACATACTGCCACAAACCTTGCATCAGGCCATGCGCTATTCCGTATTGAACGGCGGC 135

Qy         43 LysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGlu 62
              ||||||  ||||||:::|||         |||      :::|||
Db        136 AAACGCACCCGGCCCTTGTTGACTTATGCGACCGGTCAGGCTTTGGGCTTGCCGGAAAAC 195

Qy         63 LeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAsp 82
              :::           |||   |||:::|||   ||||||   ||||||||||||||||||
Db        196 GTGCTGGATGCGCCGGCTTGCGCGGTAGAATTCATCCATGTGTATTCGCTGATTCACGAC 255

Qy         83 AspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysVal 102
              ||||||||||||||||||||||||||   ||||||||||||   |||   ||||||
Db        256 GATCTGCCGGCCATGGACAACGATGATCTGCGCCGCGGCAAACCGACCTGTCACAAGGCT 315

Qy        103 TyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeu 122
              |||   |||   ||||||||||||||||||||||||||||         |||||||||:::
Db        316 TACGACGAGGCCACCGCCATTTTGGCCGGCGACGCACTGCAGGCGCTGGCCTTTGAAGTT 375

Qy        123 IleSerSerAspAspArgLeuThrAspGlu------ValLysIleLysValLeuGlnArg 140
              ::::::::::|||      :::|||   :::         :::::::||::::::
Db        376 CTGGCCAACGACCCCGGCATCACCGTCGATGCCCCGGCTCGCCTGAAAATGATCACGGCT 435

Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
              |||:::   |||||||||      ||||||||||||||||||   :::|||:::   |||
Db        436 TTGACCCGCGCCAGCGGCTCTCAAGGCATGGTGGGCGGTCAAGCCATCGATCTCGGCTCC 495

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
              |||:::   :::   |||      ||||||   ::::||      ||||||||||||
Db        496 GTCGGCCGCAAATTGACGCTGCCGGAACTCGAAAACATGCATATCCACAAGACTGGCGCC 555

Qy        181 LeuLeuThrPheAlaVal---MetSerAlaAlaAspIleAlaAsnValAspAspThrThr 199
              |||:::      :::|||   :::::::|||            :::::::|||
Db        556 CTGATCCGCGCCAGCGTCAATCTGGCGGCATTATCCAAACCCGATCTGGATACTTGCGTC 615

Qy        200 LysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAsp 219
              :::   |||:::   |||:::      :::|||:::   ||||||:::|||||||||
Db        616 GCCAAGAAACTGGATCACTATGCCAAATGCATAGGCTTGTCGTTCCAGGTCAAAGACGAC 675

Qy        220 LeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGlu 239
              :::||||||         |||   |||   |||||||||      |||   |||::::::
Db        676 ATTCTCGACATCGAAGCCGACACCGCGACACTCGGCAAGACTCAGGGCAAGGACATCGAT 735

Qy        240 AsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThr 259
              |||:::|||   ||||||   :::|||||||||      ||||||:::   |||
Db        736 AACGACAAACCGACCTACCCTGCGCTATTGGGCATGGCTGGCGCCAAACAAAAAGCCCAG 795

Qy        260 TyrHisArgAspAlaAlaValAspGluLeuThr 270
              :::   |||||||:::   ||||||
Db        796 GAATTGCACGAACAAGCAGTCGAAAGCTTAACG 828

RESULT 14
US-09-934-903-13
; Sequence 13, Application US/09934903
; Patent No. US20020102690A1
; GENERAL INFORMATION:
;  APPLICANT: Koffas, Mattheos
;  APPLICANT:  Odom, J. Martin
;  APPLICANT:  Schenzle, Andreas J.
;  APPLICANT:  No. US20020102690Alton, Kelley C.
;  APPLICANT:  Tomb, Jean-Francois
;  APPLICANT:  Rouviere, Pierre
;  APPLICANT:  Picataggio, Stephen
;  APPLICANT:  Cheng, Qiong
;  TITLE OF INVENTION: Genes Involved in Isoprenoid Compounds Production
;  FILE REFERENCE: CL1646 US NA
;  CURRENT APPLICATION NUMBER: US/09/934,903
;  CURRENT FILING DATE:  2001-08-22
;  PRIOR APPLICATION NUMBER: 60/229,907
;  PRIOR FILING DATE: September 1, 2001
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
;   LENGTH: 891
;   TYPE: DNA
;   ORGANISM: Methylomonas 16a
;   FEATURE:
;   OTHER INFORMATION: ORF7
US-09-934-903-13

Alignment Scores:
Pred. No.:              2.48e-51      Length:          891
Score:                  516.50        Matches:         113
Percent Similarity:     63.75%        Conservative:    47
Best Local Similarity:  45.02%        Mismatches:      88
Query Match:            35.55%        Indels:          3
DB:                     10            Gaps:            2

US-09-925-637-64 (1-287) x US-09-934-903-13 (1-891)

Qy         23 LysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGly 42
              ::::::::::::      |||   ::::::|||   |||||||::      ||||||
Db         76 GAAAACATACTGCCACAAACCTTGCATCAGGCCATGCGCTATTCCGTATTGAACGGCGGC 135

Qy         43 LysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGlu 62
              ||||||  ||||||:::|||         |||      :::|||
Db        136 AAACGCACCCGGCCCTTGTTGACTTATGCGACCGGTCAGGCTTTGGGCTTGCCGGAAAAC 195
```

```
Qy         63 LeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAsp 82
              :::           |||   |||:::|||   ||||||   ||||||||||||||||||
Db        196 GTGCTGGATGCGCCGGCTTGCGCGGTAGAATTCATCCATGTGTATTCGCTGATTCACGAC 255

Qy         83 AspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysVal 102
              ||||||||||||||||||||||||||   ||||||||||||   |||   ||||||
Db        256 GATCTGCCGGCCATGGACAACGATGATCTGCGCCGCGGCAAACCGACCTGTCACAAGGCT 315

Qy        103 TyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeu 122
              |||   |||   ||||||||||||||||||||||||||||         ||||||||:::
Db        316 TACGACGAGGCCACCGCCATTTTGGCCGGCGACGCACTGCAGGCGCTGGCCTTTGAAGTT 375

Qy        123 IleSerSerAspAspArgLeuThrAspGlu------ValLysIleLysValLeuGlnArg 140
              ::::::::::|||      :::|||   :::         ::::::|||:::::
Db        376 CTGGCCAACGACCCCGGCATCACCGTCGATGCCCCGGCTCGCCTGAAAATGATCACGGCT 435

Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
              |||:::   |||||||||      ||||||||||||||||||   :::|||:::   |||
Db        436 TTGACCCGCGCCAGCGGCTCTCAAGGCATGGTGGGCGGTCAAGCCATCGATCTCGGCTCC 495

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
              |||:::   :::   |||      ||||||   :::|||      ||||||||||||
Db        496 GTCGGCCGCAAATTGACGCTGCCGGAACTCGAAAACATGCATATCCACAAGACTGGCGCC 555

Qy        181 LeuLeuThrPheAlaVal---MetSerAlaAlaAspIleAlaAsnValAspAspThrThr 199
              |||:::      :::|||   :::::::|||         :::::::|||
Db        556 CTGATCCGCGCCAGCGTCAATCTGGCGGCATTATCCAAACCCGATCTGGATACTTGCGTC 615

Qy        200 LysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAsp 219
              :::   |||:::   |||:::      :::|||:::   |||||||:::|||||||||
Db        616 GCCAAGAAACTGGATCACTATGCCAAATGCATAGGCTTGTCGTTCCAGGTCAAAGACGAC 675

Qy        220 LeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGlu 239
              :::||||||         |||   |||   |||||||||      |||   |||::::::
Db        676 ATTCTCGACATCGAAGCCGACACCGCGACACTCGGCAAGACTCAGGGCAAGGACATCGAT 735

Qy        240 AsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThr 259
              |||:::|||   ||||||   :::||||||||||      |||||||:::   |||
Db        736 AACGACAAACCGACCTACCCTGCGCTATTGGGCATGGCTGGCGCCAAACAAAAAGCCCAG 795

Qy        260 TyrHisArgAspAlaAlaValAspGluLeuThr 270
                  :::   |||||||:::   ||||||
Db       .796 GAATTGCACGAACAAGCAGTCGAAAGCTTAACG 828

RESULT 15
US-09-934-868-71
; Sequence 71, Application US/09934868
; Patent No. US20020137190A1
; GENERAL INFORMATION:
;   APPLICANT: Koffas, Mattheos
;   APPLICANT:  Odom, James M
;   APPLICANT:  Schenzle, Andreas J
;   TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
;   FILE REFERENCE: CL1596 US NA
;   CURRENT APPLICATION NUMBER: US/09/934,868
;   CURRENT FILING DATE:  2001-08-22
;   PRIOR APPLICATION NUMBER: 60/229,858
;   PRIOR FILING DATE: 2000-09-01
;   NUMBER OF SEQ ID NOS: 81
;   SOFTWARE: Microsoft Office 97
; SEQ ID NO 71
;   LENGTH: 891
;   TYPE: DNA
;   ORGANISM: Methylomonas 16a
;   FEATURE:
;   OTHER INFORMATION: ORF7 ISPa
US-09-934-868-71

Alignment Scores:
Pred. No.:              2.48e-51        Length:         891
Score:                  516.50          Matches:        113
Percent Similarity:     63.75%          Conservative:   47

Best Local Similarity:  45.02%          Mismatches:     88
Query Match:            35.55%          Indels:         3
DB:                     10              Gaps:           2

US-09-925-637-64 (1-287) x US-09-934-868-71 (1-891)

Qy         23 LysSerValMetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGly 42
              ::::::::::::          |||   ::::::|||   ||||||:::      ||||||
Db         76 GAAAACATACTGCCACAAACCTTGCATCAGGCCATGCGCTATTCCGTATTGAACGGCGGC 135

Qy         43 LysArgIleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGlu 62
              ||||||   |||||||:::|||          |||      :::|||
Db        136 AAACGCACCCGGCCCTTGTTGACTTATGCGACCGGTCAGGCTTTGGGCTTGCCGGAAAAC 195

Qy         63 LeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAsp 82
              :::           |||   |||:::|||   ||||||   ||||||||||||||||||
Db        196 GTGCTGGATGCGCCGGCTTGCGCGGTAGAATTCATCCATGTGTATTCGCTGATTCACGAC 255

Qy         83 AspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysVal 102
              ||||||||||||||||||||||||||   ||||||||||||   |||   ||||||
Db        256 GATCTGCCGGCCATGGACAACGATGATCTGCGCCGCGGCAAACCGACCTGTCACAAGGCT 315

Qy        103 TyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeu 122
              |||   |||   ||||||||||||||||||||||||||||         ||||||||:::
Db        316 TACGACGAGGCCACCGCCATTTTGGCCGGCGACGCACTGCAGGCGCTGGCCTTTGAAGTT 375

Qy        123 IleSerSerAspAspArgLeuThrAspGlu------ValLysIleLysValLeuGlnArg 140
              ::::::::::|||      :::|||   :::         ::::::|||:::::
Db        376 CTGGCCAACGACCCCGGCATCACCGTCGATGCCCCGGCTCGCCTGAAAATGATCACGGCT 435

Qy        141 LeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSer 160
              |||:::   |||||||||      ||||||||||||||||||   :::|||:::   |||
Db        436 TTGACCCGCGCCAGCGGCTCTCAAGGCATGGTGGGCGGTCAAGCCATCGATCTCGGCTCC 495

Qy        161 GluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAla 180
              |||:::   :::   |||      ||||||   :::|||      ||||||||||||
Db        496 GTCGGCCGCAAATTGACGCTGCCGGAACTCGAAAACATGCATATCCACAAGACTGGCGCC 555

Qy        181 LeuLeuThrPheAlaVal---MetSerAlaAlaAspIleAlaAsnValAspAspThrThr 199
              |||:::      :::|||   :::::::|||         :::::::|||
Db        556 CTGATCCGCGCCAGCGTCAATCTGGCGGCATTATCCAAACCCGATCTGGATACTTGCGTC 615

Qy        200 LysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAsp 219
              :::   |||:::   |||:::      :::|||:::   |||||||:::|||||||||
Db        616 GCCAAGAAACTGGATCACTATGCCAAATGCATAGGCTTGTCGTTCCAGGTCAAAGACGAC 675

Qy        220 LeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGlu 239
              :::||||||         |||   |||   |||||||||      |||   |||::::::
Db        676 ATTCTCGACATCGAAGCCGACACCGCGACACTCGGCAAGACTCAGGGCAAGGACATCGAT 735

Qy        240 AsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThr 259
              |||:::|||   ||||||   :::||||||||||      |||||||:::   |||
Db        736 AACGACAAACCGACCTACCCTGCGCTATTGGGCATGGCTGGCGCCAAACAAAAAGCCCAG 795

Qy        260 TyrHisArgAspAlaAlaValAspGluLeuThr 270
                  :::   |||||||:::   ||||||
Db        796 GAATTGCACGAACAAGCAGTCGAAAGCTTAACG 828

Search completed: May 30, 2003, 22:36:09
Job time : 192 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 30, 2003, 20:17:41 ; Search time 1557 Seconds
(without alignments)
2985.295 Million cell updates/sec

Title: US-09-925-637-64
Perfect score: 1453
Sequence: 1 MTNLPMNKLIDEVNNELSVA.....ELTQIDEQFNTKHLLEIVDL 287

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09925637/runat_23052003_174818_23437/app_query.fasta_1.455
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09925637_@CGN_1_1_1906_@runat_23052003_174818_23437 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		533.5	36.7	4484	17	BH770981	BH770981 LLMGtag70
2		436.5	30.0	682	14	BQ856404	BQ856404 QGB3g04.y
3		431	29.7	685	13	BM449136	BM449136 DSA032E03
4		418	28.8	703	17	BH429531	BH429531 BOHLB93TR
5		415	28.6	709	17	BH502634	BH502634 BOHJK46TR
6		378.5	26.0	797	17	BH466685	BH466685 BOGWJ16TR
7		378.5	26.0	863	17	BH584613	BH584613 BOGQH60TR
8	c	356.5	24.5	475	17	BH439236	BH439236 BOHAC18TF
9		343	23.6	837	12	BF261534	BF261534 HV_CEa000
10	c	339.5	23.4	770	17	BH250143	BH250143 BOGAG35TF
11		336.5	23.2	442	9	AJ302129	AJ302129 AJ302129
12	c	325.5	22.4	808	14	BQ506777	BQ506777 EST614192
13		314	21.6	596	17	AQ963747	AQ963747 LERGP06TF
14	c	314	21.6	838	17	BH657017	BH657017 BOMBM36TF
15		299	20.6	666	13	BI717621	BI717621 1031021A0
16		299	20.6	675	13	BI723806	BI723806 1031067H0
17		299	20.6	701	10	BE238154	BE238154 894040E07
18		298.5	20.5	657	9	AJ502290	AJ502290 AJ502290
19		297.5	20.5	425	10	BE323612	BE323612 NF006A09P
20		297.5	20.5	558	9	AL749888	AL749888 AL749888
21		297.5	20.5	828	17	BH501300	BH501300 BOHAP61TF
22		297	20.4	384	10	BE521431	BE521431 M19E5STM
23		296	20.4	669	17	BH731669	BH731669 BOHYD34TF
24		292	20.1	668	13	BI726915	BI726915 1031088E0
25		291.5	20.1	571	10	BE432698	BE432698 EST399227
26		291.5	20.1	603	9	AI855654	AI855654 sc32d02.y
27		288.5	19.9	559	14	BQ864712	BQ864712 QGC27H09.
28		288.5	19.9	680	14	BQ983434	BQ983434 QGE19B22.
29		288.5	19.9	712	14	BU007746	BU007746 QGH5F18.y
30		288.5	19.9	720	14	BQ984560	BQ984560 QGE2h02.y
31	c	288	19.8	498	17	AQ963748	AQ963748 LERGP06TR
32		286	19.7	576	13	BI944945	BI944945 sai24d02.
33	c	284.5	19.6	771	17	BH479446	BH479446 BOGTW36TF
34		283.5	19.5	468	12	BF519441	BF519441 EST456904
35		282.5	19.4	675	14	BQ983788	BQ983788 QGE1c08.y
36		280	19.3	561	12	BF054034	BF054034 EST439264
37		280	19.3	636	9	AU238772	AU238772 AU238772
38		279	19.2	402	10	AW719932	AW719932 LjNEST12b
39		278	19.1	698	13	BI946026	BI946026 st91f10.y
40		274	18.9	722	14	BQ916912	BQ916912 QHB19O09.
41		272	18.7	355	9	AI780894	AI780894 EST261773
42		271	18.7	735	14	BQ847849	BQ847849 QGA5I17.y
43		269	18.5	628	14	BQ982915	BQ982915 QGE17L12.
44		269	18.5	716	14	BQ984280	BQ984280 QGE21D23.
45		269	18.5	745	14	BQ991706	BQ991706 QGF23J13.

ALIGNMENTS

RESULT 1
BH770981
LOCUS BH770981 4484 bp DNA linear GSS 01-MAY-2002
DEFINITION LLMGtag706 MG1363 Random Sequence Tag Library Lactococcus lactis subsp. cremoris genomic, DNA sequence.
ACCESSION BH770981
VERSION BH770981.1 GI:20373938
KEYWORDS GSS.
SOURCE Lactococcus lactis subsp. cremoris.
ORGANISM Lactococcus lactis subsp. cremoris
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.
REFERENCE 1 (bases 1 to 4484)
AUTHORS Bolotin,A., Ehrlich,S.D. and Sorokin,A.
TITLE Studies of genomes of dairy bacteria Lactococcus lactis

```
  JOURNAL   Sci. Aliments, (2002) In press
COMMENT     Contact: Sorokin A
            Genetique Microbienne
            INRA
            CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
            Tel: 33 1 34 65 25 16
            Fax: 33 1 34 65 25 21
            Email: sorokine@jouy.inra.fr
            best homologue in strain IL1403 is xseA (94%)
            Class: shotgun
            High quality sequence start: 30
            High quality sequence stop: 4456.
FEATURES             Location/Qualifiers
     source          1. .4484
                     /organism="Lactococcus lactis subsp. cremoris"
                     /strain="MG1363"
                     /db_xref="taxon:1359"
                     /clone_lib="MG1363 Random Sequence Tag Library"
                     /note="Vector: pSGMU2; Site_1: SmaI; Library of
                     chromosomal fragments of L.lactis strain MG1363 was
                     prepared by partial AluI digestion or by sonication."
BASE COUNT     1540 a    637 c   1023 g   1284 t
ORIGIN

Alignment Scores:
Pred. No.:                2.08e-56      Length:          4484
Score:                    533.50        Matches:         125
Percent Similarity:       59.34%        Conservative:    37
Best Local Similarity:    45.79%        Mismatches:      74
Query Match:              36.72%        Indels:          37
DB:                       17            Gaps:            5

US-09-925-637-64 (1-287) x BH770981 (1-4484)

Qy         25 ValMetAspThrGln---------------------------------------------- 29
              |||||||||||||:::
Db       3645 GTAATGGATACGAAGATATTAAAGCTAGAACATTTTTTAACTGAATTTTATGAAAGCGCT 3704

Qy         30 --------------LeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArg 44
                           |||   |||         |||||||||   ||||||||||||||||
Db       3705 GAATTTCCAAAAGGATTAGCTGAGCCAGCAAAATATTCCTTAATGGCGGGCGGTAAACGA 3764

Qy         45 IleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGlu------------ 60
              |||||||||:::|||   |||      |||::::::   :::   |||
Db       3765 ATTCGTCCTCTCCTTTTTTTGAATCTTTTGGAAGCTTTTGATTTGGAAGTTACAACGGCC 3824

Qy         61 ---TyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeu 79
                 |||::::::            |||   ||||||||||||||||||||   ||||||
Db       3825 CATTATCAAGTG------------GCCGCAGCGCTTGAAATGATTCATACGGGCTCGTTA 3872

Qy         80 IleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsn 99
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3873 ATTCATGATGACTTGCCTGCGATGGACAATGATGATTATCGTCGTGGAAAGTTAACGAAC 3932

Qy        100 HisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAla 119
              ||||||   :::   |||   ||||||||||||||||||||   |||
Db       3933 CATAAAAAATTTGATGAAGCGACAGCAATTCTGGCTGGAGATACTTTATTTTTCGACCCT 3992

Qy        120 PheGluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGln 139
              |||   ::::::||||||   |||   |||:::   :::       :::   :::   :::
Db       3993 TTTTATGTTCTTTCAAGTAGCGAT---TTGTCGGCGCAAACAATTGTATCATTGACCCGC 4049

Qy        140 ArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGln 159
                 |||:::   |||||||||       |||||||||   ||||||:::|||||||||
Db       4050 GAATTGGCTTATGCCTCGGGCTCGTATGGAATGGTTGCTGGTCAAATTTTGGATATGGCT 4109

Qy        160 SerGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGly 179
              ||||||:::   :::   |||      :::|||   ||||||||:::   |||||||||
Db       4110 GGTGAAGGTAAAGATTTAACTTTGACTCAGATTGAGCAAATTCATCGATTAAAAACGGGT 4169

Qy        180 AlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAspAspThrThr 199
              ||||||||||||       :::::::|||       :::|||           :::
Db       4170 CGACTTTTAACTTTTCCTTTTGTGGCTGCTGGAATCGTAGCCCAAAAAACAAATCAAGAA 4229

Qy        200 LysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAsp 219
                 |||   |||                     ||||||:::   |||||||||:::||||||
Db       4230 ATCGAAAAACTTCGTGTAGTCGGACAAATTTTAGGATTAGCTTTTCAAATTCGTGATGAC 4289

Qy        220 LeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySerAspLeuGlu 239
              :::||||||                 |||:::||||||||||      |||   |||:::
Db       4290 ATTTTAGATGTGACGGCCACTTTTGCTGAACTGGGTAAGACGCCTGGAAAGGATGTTTTA 4349

Qy        240 AsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAspLysLeuThr 259
                       ||||||||||||||:::   ||||||   :::||||||:::      ||||||
Db       4350 GAAGAAAAGTCGACTTATGTTGCGCATTTAGGACTTGAAGGAGCTAAAATATCACTGACA 4409

Qy        260 TyrHisArgAspAlaAlaValAspGluLeuThrGlnIle 272
                                  |||:::|||::::::::::
Db       4410 --------------------GATAAGTTGTCAGAAGTA 4427

RESULT 2
BQ856404
LOCUS       BQ856404                682 bp    mRNA    linear   EST 14-AUG-2002
DEFINITION  QGB3g04.yg.ab1 QG_ABCDI lettuce salinas Lactuca sativa cDNA clone
            QGB3g04, mRNA sequence.
ACCESSION   BQ856404
VERSION     BQ856404.1  GI:22241869
KEYWORDS    EST.
SOURCE      Lactuca sativa.
  ORGANISM  Lactuca sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;
            Lactuca.
REFERENCE   1  (bases 1 to 682)
  AUTHORS   Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
            Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison
            ,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
            Church,S., Jackson,L. and Bradford,K.
  TITLE     Lettuce and Sunflower ESTs from the Compositae Genome Project
            http://compgenomics.ucdavis.edu/
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Alexander Kozik [R.W.Michelmore]
            Department of Vegetable Crops, R.W.Michelmore Lab
            University of California at Davis (UCD)
            Asmundson Hall, UCD, Davis, CA 95616, USA
            Tel: 1-(530)-742-1742
            Fax: 1-(530)-752-9659
            Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
            belongs to contig QG_CA_Contig1448, see http://cgpdb.ucdavis.edu/
            for details.
            Plate: QGB3  row: g  column: 04.
FEATURES             Location/Qualifiers
     source          1. .682
                     /organism="Lactuca sativa"
                     /cultivar="Salinas"
                     /db_xref="taxon:4236"
                     /clone="QGB3g04"
                     /clone_lib="QG_ABCDI lettuce salinas"
                     /lab_host="E.coli"
                     /note="Vector: pBRcDNASfiAB; The library was constructed
                     from 10 different sources of RNA from a single genotype.
                     Separate cDNAs were generated using primers that
                     incorporated unique 5' and 3' tags to distinguish each
                     source of RNA. cDNAs were then pooled, size-fractionated,
                     directionally cloned into a custom medium-copy vector and
                     transformations made with four size classes to minimize
                     size bias. Details of each source of RNA and library
                     construction can be obtained at http://cgpdb.ucdavis.edu/
                     TAG_LIB=QG_ABCDI lettuce salinas
                     TAG_TISSUE=flowers post-fertilized
                     TAG_SEQ=TGCCATCGGG"
BASE COUNT      160 a    177 c    200 g    145 t
ORIGIN

Alignment Scores:
```

```
Pred. No.:              2.39e-45        Length:         682
Score:                  436.50          Matches:        97
Percent Similarity:     62.39%          Conservative:   44
Best Local Similarity:  42.92%          Mismatches:     76
Query Match:            30.04%          Indels:         9
DB:                     14              Gaps:           3

US-09-925-637-64 (1-287) x BQ856404 (1-682)

Qy         16 GluLeuSerValAlaIleAsnLysSerValMetAspThr---------------------- 28
              |||   :::|||   :::|||:::::::::   ::::::
Db          5 GAGAAAGCTGTTCGGGTTAACAGAGCGCTTGATGAATCCATTTCCATTAAAACGCCACCC 64

Qy         29 GlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgProVal 48
                 :::   |||:::|||   |||||||||   ||||||||||||||||:::||||||:::
Db         65 ACAATCCATGAAGCTATGCGTTACTCTCTCCTCGCCGGTGGGAAGCGTGTGAGGCCCATT 124

Qy         49 LeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSerAla 68
              |||   :::         :::   :::       :::             |||   :::|||
Db        125 CTCTGCATCGCCGCTTGCGAGCTCGTCGGAGGTGATGAATCCACCGCCATGCCAGCCGCC 184

Qy         69 IleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAsp 88
              |||:::|||||||||||||||   ||||||||||||||||||||||||   ||||||
Db        185 TGCGCTGTTGAAATGATACACACCATGTCGTTGATTCACGACGACCTTCCGTGCATGGAT 244

Qy         89 AsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrpThrAla 108
              |||||||||:::|||||||||||||   ||||||||||||||||||||||||      |||
Db        245 AACGACGATTTTCGCCGGGGGAAACCAACAAACCACAAGGTTTACGGCGAAGACGTGGCG 304

Qy        109 IleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAsp--- 127
              :::|||||||||||:::|||||||      ||||||:::   |||:::|||
Db        305 GTCCTCGCCGGCGACTCTCTCCTCGCCTTCGCTTTCCAACACATCGCAAGCGCCACCGTC 364

Qy        128 ArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHis 147
                 :::   |||      :::      :::      |||:::   :::   |||
Db        365 GGCGCGTCTCCTGAACGCGTTTTGGCCGCCGTCGGAGAACTGGCGAAGTCGATTGGGACT 424

Qy        148 ValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGly---GlnProIleAsp 166
              |||:::|||   ||||||:::::::|||:::   |||   |||   :::   :::
Db        425 GAAGGGTTGGTCGCGGGTCAAGTGGTGGACATAGCTTCAACAGGTGCAAAGGACGTCGGA 484

Qy        167 LeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaVal 186
              |||:::   ||||||      ||||||   |||||||||      ||||||
Db        485 CTGGACCAGCTTGAGTTCATCCACATCCACAAGACGGCGGCGCTATTGGAAGCAGCCGTC 544

Qy        187 MetSerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyr 206
              :::      |||   :::            ||||||      |||  |||      :::
Db        545 GTATTGGGAGCAATATTAGGTGGAGGAAGCGACACCCAAGTGGAAAAGCTGAGAAAGTTC 604

          207 SerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGlyAsp 226
              :::      :::|||::::::||||||:::   ||||||:::||||||
Db        605 GCGAGATGCATTGGATTATTGTTTCAGGTGGTTGACGACATATTAGACGTCACCAAGTCG 664

Qy        227 GluAlaLysLeuGlyLys 232
                    :::||||||||||
Db        665 TCGGAGGAATTGGGGAAA 682

RESULT 3
BM449136
LOCUS       BM449136                685 bp    mRNA    linear   EST 01-APR-2002
DEFINITION  DSA032E03_60893 An expressed sequence tag database for the
            halotolerant green alga, Dunaliella salina Dunaliella salina cDNA
            clone DSA032E03 5, mRNA sequence.
ACCESSION   BM449136
VERSION     BM449136.1  GI:19854708
KEYWORDS    EST.
SOURCE      Dunaliella salina.
  ORGANISM  Dunaliella salina
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Dunaliellaceae; Dunaliella.
REFERENCE   1  (bases 1 to 685)
  AUTHORS   Cushman,J.C.
  TITLE     An expressed sequence tag database for the halotolerant green alga,
            Dunaliella salina
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Cushman JC
            Department of Biochemistry
            University of Nevada
            MS200, Reno, NV 89557-0014, USA
            Tel: 775-784-1918
            Fax: 775-784-1650
            Email: jcushman@unr.edu
            PCR PRimers
            FORWARD: T3 20mer
            BACKWARD: T7 21mer
            Plate: 032  row: E  column: 03
            Seq primer: T3 20mer
            High quality sequence stop: 685.
FEATURES             Location/Qualifiers
     source          1. .685
                     /organism="Dunaliella salina"
                     /db_xref="taxon:3046"
                     /clone="DSA032E03"
                     /clone_lib="An expressed sequence tag database for the
                     halotolerant green alga, Dunaliella salina"
                     /tissue_type="Cells, which was adapted in 2.5M NaCI via a
                     incremental series from 1.7 to 2.0 to 2.25 to 2.5 M NaCl,
                     were exposed to 3.4 M NaCl for 5 hours"
                     /cell_type="Green"
                     /note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Library construction was performed
                     according to Stratagene's recommended protocol for the
                     Lambda UniZapXR vector and cDNA synthesis kit."
BASE COUNT      153 a    182 c    219 g    131 t
ORIGIN

Alignment Scores:
Pred. No.:              1.23e-44        Length:         685
Score:                  431.00          Matches:        92
Percent Similarity:     63.68%          Conservative:   36
Best Local Similarity:  45.77%          Mismatches:     61
Query Match:            29.66%          Indels:         12
DB:                     13              Gaps:           3

US-09-925-637-64 (1-287) x BM449136 (1-685)

Qy         78 SerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeu 97
              ||||||||||||||||||||||:::|||||||||||||||:::|||||||||:::
Db          7 AGCCTGATCCATGATGACTTGCCCTCCATGGACAACGATGACTTCCGCAGAGGGCGGCCC 66

Qy         98 ThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThr 117
              ||||||||||||||||||||||      ||||||||||||||||||||||||||:::
Db         67 ACCAACCACAAGGTGTACGGAGAGGACATTGCCATCCTTGCTGGAGATGCCCTGCTCTCT 126

Qy        118 LysAlaPheGluLeuIleSer---SerAspAspArgLeuThrAspGluValLysIleLys 136
                 |||||||||   ::::::   :::          :::   |||      ::::::
Db        127 TTCGCCTTTGAGCATGTTGCGCGCGCGACCACTGGCACCTCGCCTGAGCGTGTGCTGCGT 186

Qy        137 ValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeu 156
              |||:::      |||      |||   |||      |||:::   |||||||||::::::
Db        187 GTGATTCTGGAGCTGGGCAAGGCTGTTGGTGCAGATGGGCTGACAGGAGGACAGGTTGTG 246

Qy        157 AspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIleHisLysThr 176
              |||:::::||||||   :::   :::   ||||||   |||:::   ||||||:::
Db        247 GACATTAAGAGCGAAAACGAGGAGGTGGGCCTGGAGGTGCTGCAGTACATCCATGAACAC 306

Qy        177 LysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAsp 196
              ||||||   |||||||||      :::|||:::      |||   :::         |||
Db        307 AAGACAGCAGCCCTGCTGGAGGCATCCGTGGTGTGTGGCGCTCTCGTGGGCGGTGCGGAC 366

Qy        197 AspThrThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIle 216
              |||   |||   |||   |||      |||:::   :::::::|||:::   ||||||:::
Db        367 GACGTGACAGTGGAGAAGCTGCGCAAGTACGCGCGCAACATTGGCCTCGCGTTCCAGGTG 426

Qy        217 LysAspAspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySer 236
                 ||||||:::|||||||||                  |||||||||      |||
```

```
Db        427 GTTGACGATATCCTGGATTGCACCCAGACCACGGAGATGCTGGGCAAGACGGCAGGCAAG 486

Qy        237 AspLeuGluAsnAsnLysSerThrTyrValSerLeuLeuGly-----------LysAsp 252
              |||::::::   ||||||:::||||||      ||||||||            |||
Db        487 GACATTGATGTGAACAAGACCACGTACCCCAAACTGCTTGGCCTGGAGAAGAGTAAGCAG 546

Qy        253 GlyAlaGluAspLysLeuThrTyrHisArgAspAlaAlaValAspGluLeuThrGlnIle 272
              |||||||||   :::                    :::   :::   |||:::
Db        547 GCTGCAGAGGACCTGATT--------------------GCTGAGGCCATCCAGCAGCTG 585

Qy        273 Asp 273
              |||
Db        586 GAT 588

RESULT 4
BH429531
LOCUS       BH429531                 703 bp    DNA     linear   GSS 12-DEC-2001
DEFINITION  BOHLB93TR BOHL Brassica oleracea genomic clone BOHLB93, DNA
            sequence.
ACCESSION   BH429531
VERSION     BH429531.1  GI:17615252
KEYWORDS    GSS.
SOURCE      Brassica oleracea.
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 703)
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
  TITLE     Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished (2001)
COMMENT     Other_GSSs: BOHLB93TF
            Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .703
                     /organism="Brassica oleracea"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOHLB93"
                     /clone_lib="BOHL"
                     /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT      180 a    143 c    216 g    164 t
ORIGIN

Alignment Scores:
Pred. No.:                5.99e-43        Length:         703
Score:                    418.00          Matches:        95
Percent Similarity:       60.09%          Conservative:   39
Best Local Similarity:    42.60%          Mismatches:     85
Query Match:              28.77%          Indels:         4
DB:                       17              Gaps:           2

US-09-925-637-64 (1-287) x BH429531 (1-703)

Qy         29 GlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgProVal 48
              ::::::   |||:::|||   |||:::::::   |||   |||||||||:::||||||:::
Db         37 AAGATCCACGAAGCCATGCGTTACGCGGTTCTTGCAGCAGGGAAACGCGTTAGGCCAATA 96

Qy         49 LeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSerAla 68
              |||   |||   :::   :::   :::      |||           |||   :::|||
Db         97 CTATGCCTTGCTTCTTGCGAGCTAGTAGGAGGCGAAGAAAACGCGGCGATGCCAGCGGCT 156

Qy         69 IleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAsp 88
              |||:::||||||||||||||   |||||||||   ||||||||||||   ||||||

Db        157 TGTGCGGTTGAGATGATACACACCATGTCTCTAATCAAAGACGACTTGCCTTGTATGGAC 216

Qy         89 AsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrpThrAla 108
              |||||||||   ||||||||||||   |||   ||||||||||||||||||      |||
Db        217 AATGACGATCTGCGTCGTGGGAAGCCTACGACGCACAAAGTCTACGGCGAAGGAGTCGCT 276

Qy        109 IleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAspArg 128
              ||||||:::|||   |||||||||:::   |||||||||   :::::::::   :::
Db        277 ATTCTCTCCGGAGGAGCTCTCTTGTCTCTTGCCTTCGAGCACATGACGACGGCGGAG--- 333

Qy        129 LeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHisVal 148
              ::::::   |||      :::      ::::::   |||:::   :::   |||
Db        334 GTATCCTCGGAGAAAATGGTTTGGGCGGTCAGGGAACTGGCTAGGTCCATTGGAACTAGA 393

Qy        149 GlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIleAsp------ 166
              |||:::|||   ||||||      |||:::   |||||||||      :::|||
Db        394 GGGTTAGTCGCGGGACAAGCTAAGGACATAAGTAGTGAAGGTTTGGACTTAGACGAGGTA 453

Qy        167 ---LeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAla 185
                 ||||||   ||||||   ||||||      ||||||      ||||||      |||
Db        454 GGACTAGAGCATTTAGAGTTTATACACGTACACAAAACCGCGGTTCTTTTGGAAACTGCT 513

Qy        186 ValMetSerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSer 205
              :::      |||   |||            :::         |||   :::
Db        514 GCGGTTCTTGGAGCCATTATTGGTGGTGGGTCTAATGAAGAGGTTGAGAGAGTTAGAAAG 573

Qy        206 TyrSerTyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGly 225
              ::::::      :::|||::::::||||||:::   ||||||:::||||||
Db        574 TTCGCAAGGTGCATTGGGTTGTTGTTTCAAGTGGTGGATGATATTTTGGACGAGACCAAG 633

Qy        226 AspGluAlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyr 245
                     :::||||||||||      |||   |||           |||   ||||||
Db        634 TCGTCGGAGGAACTGGGAAAAACCGCCGGGAAAGATCAGCTCGTGGGAAAGCTGACTTAT 693

Qy        246 ValSerLeu 248
                    |||
Db        694 CCCAAGCTG 702

RESULT 5
BH502634
LOCUS       BH502634                 709 bp    DNA     linear   GSS 13-DEC-2001
DEFINITION  BOHJK46TR BOHJ Brassica oleracea genomic clone BOHJK46, DNA
            sequence.
ACCESSION   BH502634
VERSION     BH502634.1  GI:17710731
KEYWORDS    GSS.
SOURCE      Brassica oleracea.
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 709)
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
  TITLE     Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished (2001)
COMMENT     Other_GSSs: BOHJK46TF
            Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .709
                     /organism="Brassica oleracea"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOHJK46"
                     /clone_lib="BOHJ"
                     /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
```

genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT 167 a 139 c 248 g 155 t
ORIGIN

Alignment Scores:

Pred. No.:	1.48e-42	Length:	709
Score:	415.00	Matches:	91
Percent Similarity:	60.45%	Conservative:	42
Best Local Similarity:	41.36%	Mismatches:	83
Query Match:	28.56%	Indels:	4
DB:	17	Gaps:	2

US-09-925-637-64 (1-287) x BH502634 (1-709)

```
Qy         40 AlaGlyGlyLysArgIleArgProValLeuLeuLeuThrLeuAspSerLeuAsnThr 59
Db          3 GCCGGCGGCAAACGCGTCAGACCAATACTCTGCCTCGCGGCCTGCGAGCTAGTCGGAGGG 62

Qy         60 GluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeu 79
Db         63 GAAGAGCGTTCGGCGGTTCCGGCGGCTTGCGCGGTCGAGATGATTCACACGATGTCACTC 122

Qy         80 IleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsn 99
Db        123 ATCAAAGACGATCTTCCTTGCATGGACAACGACGACCTCCGGCGAGGTAAGCCCACGACG 182

Qy        100 HisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAla 119
Db        183 CACAAAGTCTTCGGAGAAAGCGTCGCGATTCTCTCCGGCGGCGCGCTCCTAGCTCTGGCC 242

Qy        120 PheGluLeuIleSerSerAspAspArgLeuThrAspGluValLysIleLysValLeuGln 139
Db        243 TTCGAGCATTTGACGGAAGCTGAC---GTGTCGCCTGAGAGAATGGTTAGAGCGGTTAAA 299

Qy        140 ArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGln 159
Db        300 GAACTGGCGAAGTCTATAGGGACGAAAGGGCTCGTGGCGGGACAAGCGATGGATTTGAGC 359

Qy        160 SerGluGly---------GlnProIleAspLeuGluThrLeuGluMetIleHisLysThr 176
Db        360 AGCGAAGGTTTGGATCAAAACGACGTCGGGTTGAAGGAGCTTGAGTTTATTCACGTTCAT 419

Qy        177 LysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAlaAsnValAsp 196
Db        420 AAAACCGGTTCGCTGCTTGAGGCTTCGGCGGTTATCGGAGCGGTTATTGGAGGTGGTTCG 479

Qy        197 AspThrThrLysGluHisLeuGluSerTyrSerTyrHisLeuGlyMetMetPheGlnIle 216
Db        480 GAGGAAGAGGTTGAGAAGGTGAGGAGGTTCGCGAGGTGTATTGGGTTGTTGTTTCAGGTG 539

          217 LysAspAspLeuLeuAspCysTyrGlyAspGluAlaLysLeuGlyLysLysValGlySer 236
Db        540 GTTGATGATATTTTGGATGGGACTAAGTCGTCGGAGGAGCTGGGGAAGACCGCCGGGAAA 599

Qy        237 AspLeuGluAsnAsnLysSerThrTyrValSerLeuLeuGlyLysAspGlyAlaGluAsp 256
Db        600 GATCAGGTGGCCGGAAAGCTGACGTATCCGAAAGTGTTGGGACTTGAGGAAAGTAAAGAG 659
```

RESULT 6
BH466685
LOCUS BH466685 797 bp DNA linear GSS 13-DEC-2001
DEFINITION BOGWJ16TR BOGW Brassica oleracea genomic clone BOGWJ16, DNA sequence.
ACCESSION BH466685
VERSION BH466685.1 GI:17667767
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 797)
AUTHORS Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other_GSSs: BOGWJ16TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
FEATURES Location/Qualifiers
source 1. .797
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOGWJ16"
/clone_lib="BOGW"
/note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT 187 a 190 c 237 g 183 t
ORIGIN

Alignment Scores:

Pred. No.:	8.79e-38	Length:	797
Score:	378.50	Matches:	92
Percent Similarity:	54.04%	Conservative:	35
Best Local Similarity:	39.15%	Mismatches:	103
Query Match:	26.05%	Indels:	5
DB:	17	Gaps:	1

US-09-925-637-64 (1-287) x BH466685 (1-797)

```
Qy         13 ValAsnAsnGluLeuSerValAlaIleAsnLysSerValMetAspThrGlnLeuGluGlu 32
Db         90 GTCAACAAGGCACTAGACGGCGCCGTTCCACTTTGCGAGCCAGTACTCAAGATCCGTGAG 149

Qy         33 SerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgProValLeuLeuLeuLeu 52
Db        150 GCCATGAGGTATACGCTTCTCTCCAACGGCAAACGCGTAAGACCAATGCTTTGTTTGGCT 209

Qy         53 ThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGlu 72
Db        210 GCATGCGAGCTCGTTGGCGGCCATGAGTCAACCGCAATGTCTTTTGCGTGTGCACTTGAG 269

Qy         73 MetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyr 92
Db        270 ATGATCCACGCGTCGTCTATCATCTTAGACGACCTCCCTTGCATGGACGACGACAGCCTC 329

Qy         93 ArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGly 112
Db        330 CGCCGCGGAAAACCCACCAACCATAAAGTATTTGGCGAGGATAGCGCCGTTTTGGCCTCT 389

Qy        113 AspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAspArgLeuThrAspGlu 132
Db        390 GTTGCGCTCATAGGTTTGGCCGTCAAACAAACCTCTTCATCCCCGGATGTTCCTCCAGAG 449

Qy        133 ValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGly 152
Db        450 AGAGTTCTCCGGGCGGTTCAGGTGATGATGCGAGCGGTGGGAACAGAAGGACTCGTTGCG 509

Qy        153 GlyGlnMetLeuAspMetGlnSerGluGlyGlnProIleAsp--------------Leu 167
Db        510 GGACAAGCGGAGGATCTTGCCGGAGAAGGGATGAGCTTCGAGGAAAATGAAGCGGCGCTG 569

Qy        168 GluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMet 187
Db        570 GAACATCTTGAGTTTATACATATTCATAAAACGGCGTCGCTGCTTGAAGCTGCTGCGGTG 629

Qy        188 SerAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyrSer 207
Db        630 GTGGGAGCCATAATGGGAGGTGGCTCTGATGAGGAGATTGAAAGGCTTGGGAAATACGCG 689
```

```
Qy        208 TyrHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGlyAspGlu 227
              :::|||:::|||||||||:::   ||||||:::|||||
Db        690 AGGTGCGTTGGGCTGATGTTTCAAGTGGTGGATGATGTGCTCGACGTGACTAAGTCGTCA 749

Qy        228 AlaLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLys 242
              :::||||||||||      |||   |||:::         |||
Db        750 GAAGAGCTGGGGAAAACCGCAGGTAAAGATGTGATGGCTGGGAAG 794

RESULT 7
BH584613
LOCUS       BH584613                863 bp    DNA     linear   GSS 15-DEC-2001
DEFINITION  BOGQH60TR BOGQ Brassica oleracea genomic clone BOGQH60, DNA
            sequence.
ACCESSION   BH584613
VERSION     BH584613.1  GI:17837070
KEYWORDS    GSS.
SOURCE      Brassica oleracea.
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 863)
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
  TITLE     Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished (2001)
COMMENT     Other_GSSs: BOGQH60TF
            Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .863
                     /organism="Brassica oleracea"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOGQH60"
                     /clone_lib="BOGQ"
                     /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT      217 a    190 c    241 g    215 t
ORIGIN

Alignment Scores:
Seed. No.:              1.01e-37        Length:         863
Score:                  378.50          Matches:        87
Percent Similarity:     55.95%          Conservative:   40
Best Local Similarity:  38.33%          Mismatches:     97
Query Match:            26.05%          Indels:         3
DB:                     17              Gaps:           2

US-09-925-637-64 (1-287) x BH584613 (1-863)

Qy         13 ValAsnAsnGluLeuSerValAlaIleAsnLysSerValMetAspThrGlnLeuGluGlu 32
              ||||||      |||      |||:::                    ::::::   |||
Db         39 GTCAACAAAGCACTATACGAGGCTGTTCCTCTCTGCGAGCCAGCACTCAAAATCCGTGAA 98

Qy         33 SerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIleArgProValLeuLeuLeuLeu 52
              ||||||   |||:::|||   :::|||||||||   :::||||||:::|||   |||
Db         99 TCCATGAGGTACACACTTCTCTCGGGTGGAAAATGTGTAAGGCCAATGCTTTGCTTAGCT 158

Qy         53 ThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGlu 72
              :::   :::      :::      :::   |||   :::|||   |||:::|||
Db        159 GCGTGCGAGCTCGTAGGCGCCCAAGAGTCTATCGCAATGCCCACTGCATGTGCGGTTGAG 218

Qy         73 MetIleHisThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyr 92
              |||:::|||         ||||||   ||||||||||||   ||||||:::|||:::
Db        219 ATGCTTCACGCGTCATTTCTCATCCAAGACGACCTCCCTTGTATGGATGACGACAACCTT 278

Qy         93 ArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGly 112
              ||||||||||||   |||||||||||||||||||||||||      ::::::|||
Db        279 CGTCGAGGAAAACCCACCAACCATAAAGTATATGGTGAAGATATATCCGTCTTGACCGCT 338

Qy        113 AspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAspArgLeu------Thr 130
              |||   |||:::      |||   :::   :::::::::          |||
Db        339 GATGGACTCATAGCTTTGGCTGTCAAGAAGATGGCCGCGTCTACCTTTTTGGGCGTTCCT 398

Qy        131 AspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMet 150
              |||      ::::::   :::      ::::::   |||   |||      |||:::
Db        399 CCAGAGAGGGTTCTCCGGGCCATTCTGGAGATGACAAAAGCGTTGGGAACGGAAGGGCTC 458

Qy        151 ValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnPro---IleAspLeuGluThr 169
              |||   ||||||      |||:::      ||||||          :::   ||||||
Db        459 GTTGCGGGTCAAGCGGCTGATCTGGCCGGAGAAGGAATGAGCGGAGTGGGGTTGGAACAT 518

Qy        170 LeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSerAla 189
              ||||||   ||||||      ||||||   ||||||||      :::   :::
Db        519 CTTGAGTTTATACATATTAACAAAACGGCGGCGTTGCTTGAAGCTTCTGCGGTGGTGGGG 578

Qy        190 AlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyrSerTyrHis 209
              |||   :::          |||          |||   |||      |||:::
Db        579 GCTATAGTGGGAGGTGGGTCTGATGAAGAGATTGAAAGGCTTAGGAAGTATGCAAGATGC 638

Qy        210 LeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGlyAspGluAlaLys 229
              :::|||:::|||||||||:::   ||||||:::||||||                 :::
Db        639 ATTGGGCTTATGTTTCAAGTTGTGGATGATGTTCTTGACGTGACAAAGTCGTCGGAGGAG 698

Qy        230 LeuGlyLysLysValGlySer 236
              |||||||||      ||||||
Db        699 TTAGGGAAGACGGCCGGTTCA 719

RESULT 8
BH439236/c
LOCUS       BH439236                475 bp    DNA     linear   GSS 12-DEC-2001
DEFINITION  BOHAC18TF BOHA Brassica oleracea genomic clone BOHAC18, DNA
            sequence.
ACCESSION   BH439236
VERSION     BH439236.1  GI:17624950
KEYWORDS    GSS.
SOURCE      Brassica oleracea.
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 475)
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
  TITLE     Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished (2001)
COMMENT     Other_GSSs: BOHAC18TR
            Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .475
                     /organism="Brassica oleracea"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOHAC18"
                     /clone_lib="BOHA"
                     /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT       97 a    163 c    122 g     93 t
ORIGIN

Alignment Scores:
```

```
Pred. No.:              2.43e-35        Length:         475
Score:                  356.50          Matches:        75
Percent Similarity:     66.44%          Conservative:   24
Best Local Similarity:  50.34%          Mismatches:     47
Query Match:            24.54%          Indels:         3
DB:                     17              Gaps:           1

US-09-925-637-64 (1-287) x BH439236 (1-475)

Qy         36 TyrSerLeuAsnAlaGlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAsp 55
              |||||||||   ||||||||||||||:::|||||||||||||   :::          :::
Db        474 TACTCTCTCCTCGCCGGCGGGAAGCGAGTGAGACCGGTTCTCTGCATCGCCGCGTGCGAG 415

Qy         56 SerLeuAsnThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHis 75
                 :::      :::      :::   :::   :::|||   |||:::|||||||||||||
Db        414 CTGGTCGGCGGCAAAGAATCCGTCGCTTTGCCGGCGGCTTGCGCGGTGGAAATGATCCAC 355

Qy         76 ThrTyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGly 95
              |||   |||||||||||||||||||||||   ||||||||||||||   |||||||||
Db        354 ACGATGTCGCTGATCCACGACGATCTCCCGTGTATGGACAACGACGACCTCCGCCGCGGA 295

Qy         96 LysLeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeu 115
              |||   ||||||||||||||:::||||||      |||:::|||||||||||||||||||
Db        294 AAACCCACCAACCACAAAGTGTTCGGCGAAGACGTCGCGGTTTTAGCCGGAGACGCGCTT 235

Qy        116 LeuThrLysAlaPheGluLeuIleSerSerAspAspArgLeuThrAspGluValLysIle 135
              |||:::   |||||||||   :::::|||       :::                 :::
Db        234 TTATCGTTCGCGTTTGAGCACTTGGCGTCGACGTCGGGGGTGGCTCCGGCGAGGGTGGTT 175

Qy        136 LysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMet 155
              :::   :::      |||:::   |||   |||      |||:::|||   ||||||:::
Db        174 AGAGCGATCGGGGAGCTGGCGAGAGCCGTTGGATCAGAAGGGCTTGTGGCGGGTCAGATT 115

Qy        156 LeuAspMetGlnSerGluGly--------GlnProIleAspLeuGluThrLeuGluMet 172
              :::|||:::   |||||||||             :::   ||||||   ||||||
Db        114 GTGGATATCAGCAGCGAAGGTATGGACTCAAACGACGTCGGATTAGAGCTACTTGAGTTT 55

Qy        173 IleHisLysThrLysThrGlyAlaLeu 181
              ||||||      ||||||      |||
Db         54 ATCCACGTTCACAAAACGGCGGTTTTG 28

RESULT 9
BF261534
LOCUS       BF261534                 837 bp    mRNA    linear   EST 23-OCT-2001
DEFINITION  HV_CEa0001I15f Hordeum vulgare seedling green leaf EST library
            HVcDNA0004 (Blumeria challenged) Hordeum vulgare cDNA clone
            HV_CEa0001I15f, mRNA sequence.
ACCESSION   BF261534
VERSION     BF261534.3  GI:16333915
KEYWORDS    EST.
SOURCE      Hordeum vulgare.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 837)
  AUTHORS   Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Wei,F., Begum,D.,
            Frisch,D., Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi
            ,D.W., Fenton,R.D., Oates,R. and Main,D.
  TITLE     Development of a genetically and physically anchored EST resource
            for barley genomics: Blumeria infected incompatible (Mla13)
            seedling leaf cDNA library
  JOURNAL   Unpublished (2001)
COMMENT     On Nov 17, 2000 this sequence version replaced gi:13259849.
            Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Total hq bases = 232
            Seq primer: AATTAACCCTCACTAAAGGG
            High quality sequence stop: 611.
FEATURES             Location/Qualifiers
     source          1. .837
                     /organism="Hordeum vulgare"
                     /cultivar="CI16155 (Mla13)"
                     /db_xref="taxon:4513"
                     /clone="HV_CEa0001I15f"
                     /clone_lib="Hordeum vulgare seedling green leaf EST
                     library HVcDNA0004 (Blumeria challenged)"
                     /tissue_type="seedling green leaf"
                     /lab_host="TJC121"
                     /note="Vector: lambdaZAP; Site_1: EcoR1; Site_2: Xho1;
                     C.I. 16155 (Mla13) plants were greenhouse grown in the R
                     Wise lab at Iowa State University, Ames, IA; 7 day old
                     green seedlings were challenged with isolate A27 (AvrMla13
                     ) of Blumeria graminis f. sp. hordei, and leaves were
                     harvested 20 and 24 hr post-inoculation and snap frozen;
                     uninoculated leaves were harvested 20 hr post-inoculation
                     (Wei, Wise). In the TJ Close lab at the University of
                     California, Riverside, total RNA was prepared from each
                     sample pool, equal quantities of all three RNA pools were
                     combined, poly(A) RNA was purified from the mixture, one
                     cDNA library was made, and 1 million pfu were in vivo
                     excised to give pBluescript SK(-) cDNA phagemids (Choi,
                     Close). Phagemids were plated and picked at the Clemson
                     University Genomics Institute (CUGI) (Begum, Palmer,
                     Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
                     sequencing and sequence analysis were performed at CUGI
                     (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).
                     The sequence has been trimmed to remove vector sequence
                     and contains a minimum of 100 bases of phred value 20 or
                     above. For more details on library preparation and
                     sequence analysis see
                     http://www.genome.clemson.edu/projects/barley. To order
                     this clone see http://www.genome.clemson.edu/orders Also
                     see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
                     Genetically and physically anchored EST resources for
                     barley genomics. Barley Genetics Newsletter 31:29-30.
                     (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"
BASE COUNT      157 a    258 c    250 g    171 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              3.47e-33        Length:         837
Score:                  343.00          Matches:        87
Percent Similarity:     56.16%          Conservative:   36
Best Local Similarity:  39.73%          Mismatches:     78
Query Match:            23.61%          Indels:         18
DB:                     12              Gaps:           5

US-09-925-637-64 (1-287) x BF261534 (1-837)

Qy         45 IleArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGly 64
              :::|||||||   |||   |||
Db          1 GTGCGCCCGGCCCTCTGCCTGGCCGCCTGCGTCGTGTCCGGCGGCCGCGAGGCCTGGGCG 60

Qy         65 MetLysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeu 84
              |||      |||   |||:::||||||:::||||||   |||||||:::||||||||||||
Db         61 ATGGCCCCCGCCGCCGCGGTCGAGATGGTGCACACCATGTCGCTCGTACACGACGACCTT 120

Qy         85 ProAlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGly 104
              |||   ||||||:::||||||   ||||||||||||   |||   |||   ::::::|||
Db        121 CCCTGCATGGACGACGACGACCTCCGTCGTGGCAAGCCCACCTGCCACGTCATGTTCGGC 180

Qy        105 GluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSer 124
              |||      |||   ||||||||||||||||||||      ||||||:::   ::::::
Db        181 GAGCCCATCGCCGCGCTCGCCGGCGACGCCCTGCTCGCGCTCGCCTTCCAGCACATGGCC 240

Qy        125 SerAspAsp--------------------------ArgLeuThrAspGluValLysIle 135
              |||   |||                        :::   ||||||   |||   :::
Db        241 AGCGTCGACTCCTACCTTGCGGACGTTGACCCCGGCAAGCACACCGACCGCGTT---GTC 297

Qy        136 LysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMet 155
```

```
Db        298 CGCGCCATTGGTGAGCTCGCGCGCTGCATCGGATCATAGGGCCTCGTTGGCGGTCAGGTT 357

Qy        156 LeuAspMetGlnSer------GluGlyGlnProIleAspLeuGluThrLeuGluMetIle 173
Db        358 GTTGATCTGGAGATGACTGTGTCAACTGAGCCTGTACCACTTGACCGCCTGGAGTACATT 417

Qy        174 HisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAlaAspIleAla 193
Db        418 CATGTGCACAAGACCGCTGCATTGCTTGAGGCCTCTCGGGTTATTGGTGCAATCATTGCG 477

Qy        194 AsnValAspAspThrThrLysGluHisLeuGluSer---------TyrSerTyrHisLeu 210
Db        478 GGTGGCTCAGAC---------GAGCACATCGAGCAGGTGCTCATGTACGCCAGATCACTC 528

Qy        211 GlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGlyAspGluAlaLysLeu 230
Db        529 GAGCTGCTCGATCGCGCGGATGACGACATCCTTGATGTCACCAAGTGATCAGATGAGCTA 588

          231 GlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyrValSerLeuLeu 249
          589 GGGAAGACAGCTGGTAACGACTTGCCGAGTGACAAGACCCCATACCCCCATTTACTG 645

RESULT 10
BH250143/c
LOCUS       BH250143                770 bp    DNA     linear   GSS 26-NOV-2001
DEFINITION  BOGAG35TF BOGA Brassica oleracea genomic clone BOGAG35, DNA
            sequence.
ACCESSION   BH250143
VERSION     BH250143.1  GI:17073585
KEYWORDS    GSS.
SOURCE      Brassica oleracea.
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 770) .
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
  TITLE     Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished (2001)
COMMENT     Other_GSSs: BOGAG35TR
            Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .770
                     /organism="Brassica oleracea"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOGAG35"
                     /clone_lib="BOGA"
                     /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT      190 a    202 c    182 g    196 t
ORIGIN

Alignment Scores:
Pred. No.:              8.47e-33        Length:         770
Score:                  339.50          Matches:        80
Percent Similarity:     62.23%          Conservative:   37
Best Local Similarity:  42.55%          Mismatches:     58
Query Match:            23.37%          Indels:         13
DB:                     17              Gaps:           5

US-09-925-637-64 (1-287) x BH250143 (1-770)

Qy          6 MetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerVal 25

Db        539 GTGAACAAAGCTCTAGACGAA--------------GCCATACCAGTCGGGGAGCCACTC 495

Qy         26 MetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIle 45
Db        494 ---------AAGATCCACGAAGCCATGCGTTACGCGGTTCTTGCAGCCGGAAAACGTGTA 444

Qy         46 ArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMet 65
Db        443 AGGCCTATTCTCTGCCTCGCTTCTTGCGAGCTAGTAGGCGGCGAAGAAAACGCCGCGATG 384

Qy         66 LysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
Db        383 CCAGCGGCTTGTGCGGTTGAGATGATACACACCATGTCTCTTATCAAAGACGACTTGCCT 324

Qy         86 AlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGlu 105
Db        323 TGCATGGACAATGACGACCTGCGTCGTGGGAAGCCCACCACTCACAAAGCCTTCGGCGAA 264

Qy        106 TrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSer 125
Db        263 GGAATCGCCATTCTCTCCGGAGGAGCTCTCTTGTCTCTTGCGTTTGAGCACATGACAACG 204

Qy        126 AspAspArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSer 145
Db        203 GCTGAG---ATATCCTCGGAGAGAATGGTTTGGGCGGTCAGGGAACTGGCTAGGTCCATT 147

Qy        146 GlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGly--------- 162
Db        146 GGAACTAGAGGGTTGGTCGCGGGACAAGCCAAGGACATAAGTAGTGAAGGTTTGGAGTTG 87

Qy        163 GlnProIleAspLeuGluThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeu 182
Db         86 AACGACGTCGGTTTAGAGCATTTGGAGTTTATCCATGTACACAAAACCGCTGTTTTGTTG 27

Qy        183 ---ThrPheAlaValMetSerAla 189
Db         26 GAAACTGCTGCGGTTCTTGGAGCC 3

RESULT 11
AJ302129
LOCUS       AJ302129                442 bp    mRNA    linear   EST 03-APR-2001
DEFINITION  AJ302129 Solanum tuberosum cv. Bintje leaf Solanum tuberosum cDNA
            clone S3-2B, mRNA sequence.
ACCESSION   AJ302129
VERSION     AJ302129.1  GI:13536952
KEYWORDS    EST.
SOURCE      potato.
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 442)
  AUTHORS   Beyer,K., Binder,A., Boller,T. and Collinge,M.A.
  TITLE     Identification of potato genes induced during colonization by
            Phytophthora infestans
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Beyer K
            Plant Biology
            Friedrich Miescher-Institut
            Maulbeerstrasse 66, 4058 Basel, Switzerland.
FEATURES             Location/Qualifiers
     source          1. .442
                     /organism="Solanum tuberosum"
                     /cultivar="Bintje"
                     /db_xref="taxon:4113"
                     /clone="S3-2B"
                     /clone_lib="Solanum tuberosum cv. Bintje leaf"
                     /tissue_type="leaf"
                     /note="infected with Phytophthora infestans"
BASE COUNT      104 a    102 c    125 g    111 t
ORIGIN
```

```
Alignment Scores:
Pred. No.:              7.96e-33        Length:         442
Score:                  336.50          Matches:        67
Percent Similarity:     69.57%          Conservative:   29
Best Local Similarity:  48.55%          Mismatches:     41
Query Match:            23.16%          Indels:         1
DB:                     9               Gaps:           1

US-09-925-637-64 (1-287) x AJ302129 (1-442)

Qy         37 SerLeuAsnAlaGlyGlyLysArgIleArgProValLeuLeuLeuLeuThrLeuAspSer 56
              ||||||   |||||||||||||||||:::|||||||:::|||   :::   :::   :::
Db          1 TCACTTCTAGCTGGAGGAAAACGTGTCCGGCCGATCCTCTGCATGGCTTCTTGTGAAGTT 60

Qy         57 LeuAsnThrGluTyrGluLeuGlyMetLysSerAlaIleAlaLeuGluMetIleHisThr 76
              :::      :::      |||   :::   :::|||   |||:::|||||||||||||||
Db         61 GTAGGAGGGGATGAATCCTTAGCTATACCTGCAGCTTGCGCGGTTGAGATGATCCATACC 120

Qy         77 TyrSerLeuIleHisAspAspLeuProAlaMetAspAsnAspAspTyrArgArgGlyLys 96
                 ||||||:::|||||||||||||||   |||||||||||||||   ||||||||||||
Db        121 ATGTCACTCGTCCATGATGATCTTCCTTGCATGGACAACGATGATCTACGTCGTGGCAAG 180

Qy         97 LeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGlyAspAlaLeuLeu 116
              ||||||||||||||||:::||||||   ||||||:::|||   ||||||||||||||||
Db        181 CCCACGAACCATAAGGTTTTTGGGGAAAACACTGCAGTTCTTGTAGGGGATGCACTTTTG 240

Qy        117 ThrLysAlaPheGluLeuIleSerSer---AspAspArgLeuThrAspGluValLysIle 135
              :::   ||||||||   :::::::::           :::|||   :::      :::
Db        241 TCTTTGGCCTTTGAACATGTGGCTACCAAGACTCAGAATGTGACACCCCAAAGAGTGGTT 300

Qy        136 LysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMetValGlyGlyGlnMet 155
              :::   :::      |||      |||   |||      |||:::|||   ||||||:::
Db        301 CAAGCCATTGGGGAATTGGGTTCAGCCGTTGGCTCAGAAGGGCTCGTGGCAGGGCAGATT 360

Qy        156 LeuAspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeuGluMetIle 173
              :::|||:::   |||||||||:::   :::   |||      ||||||   :::
Db        361 GTGGACTTGGCGAGTGAGGGAAAACAAGTTAGCCTAACTGAACTGGAGTACCTC 414

RESULT 12
BQ506777/c
LOCUS       BQ506777                808 bp    mRNA    linear   EST 22-JUL-2002
DEFINITION  EST614192 Generation of a set of potato cDNA clones for microarray
            analyses mixed potato tissues Solanum tuberosum cDNA clone STMGM34
            3' end, mRNA sequence.
ACCESSION   BQ506777
VERSION     BQ506777.2  GI:21922630
KEYWORDS    EST.
SOURCE      potato.
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 808)
  AUTHORS   Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C.,
            Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and
            Karamycheva,S.A.
  TITLE     Generation of a set of potato cDNA clones for microarray analyses
  JOURNAL   Unpublished (2002)
COMMENT     On Jun 10, 2002 this sequence version replaced gi:21365646.
            Other_ESTs: EST614191
            Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato@tigr.org
            This clone is available through the Research Genetics, contact the
            Research Genetics for further information 1-800-711-6195 or
            cdna@resgen.com
            Seq primer: T7.
FEATURES             Location/Qualifiers
     source          1. .808
                     /organism="Solanum tuberosum"
                     /cultivar="Kennebec or Binjte"
                     /db_xref="taxon:4113"
                     /clone="STMGM34"
                     /clone_lib="Generation of a set of potato cDNA clones for
                     microarray analyses mixed potato tissues"
                     /tissue_type="mixed tissues"
                     /lab_host="SOLR"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; supplier: Combination of untreated and Phytophthora
                     infestans-treated libraries of stolons, leaves, leaflets,
                     axillary buds of stem explants, petioles, germinating eyes
                     , tubers, or roots."
BASE COUNT      238 a    208 c    143 g    219 t
ORIGIN

Alignment Scores:
Pred. No.:              5.77e-31        Length:         808
Score:                  325.50          Matches:        78
Percent Similarity:     54.34%          Conservative:   41
Best Local Similarity:  35.62%          Mismatches:     96
Query Match:            22.40%          Indels:         4
DB:                     14              Gaps:           3

US-09-925-637-64 (1-287) x BQ506777 (1-808)

Qy         71 LeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeu-ProAlaMetAspAsnAs 90
              :::||||||:::|||          |||:::|||||||||||| |||   ||||||||||
Db        793 ATGGAGATGATGCATGCTATGTGGTTGATGCACGACGACCTTTCCCTGCATGGACAATGA 734

Qy         90 pAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLe 110
              ||||:::||||||||||||||:::|||||||||||||||||||||          :::||
Db        733 TGATTTCCGTCGAGGAAAGCTGTCAAATCACAAGGTTTATGGCGAAAACGTCACTGTTTT 674

Qy        110 uAlaGlyAspAlaLeuLeuThrLysAlaPheGlu---LeuIleSerSerAspAspArgLe 129
              |||||||   :::|||:::      ||||||:::   :::   ::::::          ::
Db        673 AGCTGGTTATTCCCTTGTTGCCTTTGCATTTCAGCATATGACAACAGCCACTAAAGGGGT 614

Qy        129 uThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGl 149
              :      :::      ::::::   :::      |||:::          |||      ||
Db        613 ACACCCGAAAACAATGGTTCGTGCTGTTGGAGAACTAGCAAGATTGATAGGTCCAGAGGG 554

Qy        149 yMetValGlyGlyGlnMetLeuAspMetGlnSerGluGly---GlnProIleAspLeuGl 168
              |:::|||   ||||||:::||||||:::          |||                |||||
Db        553 GGTAGTGGCTGGCCAGGTGCTTGACTTGCTGTGTGGAGGTAACTCTGATACTGGATTAGA 494

Qy        168 uThrLeuGluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSe 188
              |   ||||||   ||||||:::   ||||||                    |||   :::
Db        493 AGAGCTTGAGTATATTCATCGTCACAAGACTGCAGACTTTGCAGAGGCCGCGGCCATTGT 434

Qy        188 rAlaAlaAspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyrSerTy 208
                 |||   :::          :::                    |||      :::|||
Db        433 AGGAGCATTGGTTGGTGGTGCGTCTGAGGAGGAGATTAATAGACTTAGGAAATTCTCCCA 374

Qy        208 rHisLeuGlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGlyAspGluAl 228
                 |||:::|||:::|||:::   ||||||:::||||||
Db        373 GTGTTTTGGACTGATGTATCAGGTTGTGGATGATATTCTTGATGTCACTAAATCCTCTGA 314

Qy        228 aLysLeuGlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyrValSerLe 248
               :::||||||||||      |||   ||||||      ||||||   ||||||      ::
Db        313 GCAATTAGGAAAGACAGCGGGGAAGGATTTGTTGGCTAACAAATTGACGTACCCAAAGAT 254

Qy        248 uLeuGlyLysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAlaValAspGl 268
              ::::|||   |||   :::::::::                          |||   :::::
Db        253 GATTGGCATTGACAAGTCCAAAGAATATGCTCAAAAACTTAACAAGGAGGCCAAGGAGCA 194

Qy        268 uLeuThrGlnIleAspGluGlnPheAsnThrLysHisLeuLeuGluIleValAsp 286
              :|||        |||   :::      :::      ||||||   :::   |||
Db        193 ACTTGTTGGTTTTGATCCAAAA---AAGTCAGCTCCATTACTTGCTTTGGCAGAT 142

RESULT 13
AQ963747
LOCUS       AQ963747                596 bp    DNA     linear   GSS 28-JAN-2000
DEFINITION  LERGP06TF LERG Arabidopsis thaliana genomic clone LERGP06, DNA
            sequence.
ACCESSION   AQ963747
```

```
VERSION     AQ963747.1  GI:6791448
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 596)
  AUTHORS   Buell,C.R., Lin,X., Pai,G., Barnstead,M., Bowman,C., Utterbach,T.,
            Feldblyum,T., Liang,F., Creasy,T. and Fraser,C.M.
  TITLE     Genomic survey sequencing of Landsberg erecta ecotype of
            Arabidopsis thaliana and identification of sequence-based
            polymorphisms
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Xiaoying Lin
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: at@tigr.org
            For addtional information, see http://www.tigr.org/tdb/at/at.html
            Seq primer: TF
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .596
                     /organism="Arabidopsis thaliana"
                     /strain="Landsberg erecta"
                     /db_xref="taxon:3702"
                     /clone="LERGP06"
                     /clone_lib="LERG"
                     /note="Organ: Leaf; Vector: pUC19JK; Total genomic DNA was
                     sheared to 0.4-0.7 Kbp before ligation."
BASE COUNT       88 a    218 c    208 g     82 t
ORIGIN

Alignment Scores:
Pred. No.:              1.03e-29        Length:         596
Score:                  314.00          Matches:        71
Percent Similarity:     56.76%          Conservative:   34
Best Local Similarity:  38.38%          Mismatches:     78
Query Match:            21.61%          Indels:         2
DB:                     17              Gaps:           1

US-09-925-637-64 (1-287) x AQ963747 (1-596)

Qy         93 ArgArgGlyLysLeuThrAsnHisLysValTyrGlyGluTrpThrAlaIleLeuAlaGly 112
              |||||||||:::   |||   |||   |||:::||||||      ||||||||||||||
Db          1 CGCCGCGGCCGCCCCACCTGCCACATCGTGTTCGGCGAGGCGATGGCCATCCTGGCTGGC 60

Qy        113 AspAlaLeuLeuThrLysAlaPheGluLeuIleSerSerAspAsp------ArgLeuThr 130
              |||||||||         |||||||||:::::::::   |||         :::   :::
Db         61 GATGCGCTGCAGGCGCTGGCGTTCGAAATTCTGGCCGAGGACACCACCCGCCAGACCAGC 120

Qy        131 AspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSerGlyHisValGlyMet 150
                 |||      :::   :::|||:::   |||      |||   |||      ||||||
Db        121 CCCGAGACCGGCATGGCCATGCTGCGCACGCTGGGCCGCGCCTGCGGTGCCGAAGGCATG 180

Qy        151 ValGlyGlyGlnMetLeuAspMetGlnSerGluGlyGlnProIleAspLeuGluThrLeu 170
              |||||||||   ||||||:::   :::   |||||||||:::   |||      |||
Db        181 GCCGGCGGCCAGGCGCTGGACCTTGCCGCCGTGGGCCAGCCGCTGACGCTGGCCCAGCTG 240

Qy        171 GluMetIleHisLysThrLysThrGlyAlaLeuLeuThrPheAlaValMetSerAlaAla 190
              |||   :::|||      |||||||||||||||:::      :::|||          |||
Db        241 GAGCACATGCATGCCTGCAAGACCGGCGCGCTGATCCGCGCCTCAGTGCAACTGGGTGCG 300

Qy        191 AspIleAlaAsnValAspAspThrThrLysGluHisLeuGluSerTyrSerTyrHisLeu 210
                 |||      |||                 |||:::   |||:::         :::
Db        301 CTGGCCGCGGGCGGCGATGCCGACGAGCTGCTGGCACTGGATGAATACGCCAGCGCCGTG 360

Qy        211 GlyMetMetPheGlnIleLysAspAspLeuLeuAspCysTyrGlyAspGluAlaLysLeu 230
              |||:::   ||||||:::::|||||||:::|||||||      |||:::         :::
Db        361 GGCCTGGCGTTCCAGGTGCGCGACGACATCCTGGACGTTGAAGGCGAATCCACCGTGATC 420

Qy        231 GlyLysLysValGlySerAspLeuGluAsnAsnLysSerThrTyrValSerLeuLeuGly 250
              ||||||      |||   |||             |||   |||:::   |||:::::|||
Db        421 GGCAAGACCGCCGGCAAGGACGCGGCCGCCGGCAAACCCACCTTCCCCTCGATCATCGGC 480

Qy        251 LysAspGlyAlaGluAspLysLeuThrTyrHisArgAspAlaAlaValAspGluLeuThr 270
                 |||   :::      :::|||             |||   |||:::      :::
Db        481 CTCGACGCCTCGCGCGCACGGCTGGCCGAGCTGACCGACGTCGCACTGCATGCCATCGCG 540

Qy        271 GlnIleAspGluGln 275
                 :::   |||:::
Db        541 CCGCTGGGCGAACGC 555

RESULT 14
BH657017/c
LOCUS       BH657017                838 bp    DNA     linear   GSS 19-FEB-2002
DEFINITION  BOMBM36TF BO_2_3_KB Brassica oleracea genomic clone BOMBM36, DNA
            sequence.
ACCESSION   BH657017
VERSION     BH657017.1  GI:18715378
KEYWORDS    GSS.
SOURCE      Brassica oleracea.
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 838)
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
  TITLE     Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .838
                     /organism="Brassica oleracea"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOMBM36"
                     /clone_lib="BO_2_3_KB"
                     /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT      179 a    200 c    231 g    228 t
ORIGIN

Alignment Scores:
Pred. No.:              1.84e-29        Length:         838
Score:                  314.00          Matches:        69
Percent Similarity:     62.42%          Conservative:   29
Best Local Similarity:  43.95%          Mismatches:     57
Query Match:            21.61%          Indels:         2
DB:                     17              Gaps:           2

US-09-925-637-64 (1-287) x BH657017 (1-838)

Qy          6 MetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerVal 25
              |||             :::   :::|||      |||      ||||||         |||
Db        487 ATGAAAACCAAAGCAGAGTCAATAAACAGAGCGCTAGACAAAGCCATC---CCACTCGTC 431

Qy         26 MetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIle 45
                    :::   :::::::|||   |||::::::   ||||||||||||||:::
Db        430 GAGCCTCTCAACATCCACAAGGCAATGCGCTACGCGATCCTCGCCGGCGGCAAACGCGTC 371

Qy         46 ArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMet 65
              ||||||:::|||   |||         :::   :::      |||         :::
Db        370 AGACCAATACTCTGCCTCGCGGCCTGCGAGCTAGTCGGAGGGGAAGAGCGTTCGGCGGTT 311

Qy         66 LysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
```

```
                :::|||   |||:::|||||||||||||||   |||||||||   ||||||||||||
Db        310 CCGGCGGCTTGCGCGGTCGAGATGATTCACACGATGTCACTCATCAAAGACGATCTTCCT 251

Qy         86 AlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGlu 105
              ||||||||||||||   ||||||||||||   |||   |||||||||:::||||||
Db        250 TGCATGGACAACGACGACCTCCGGCGAGGTAAGCCCACGACGCACAAAGTCTTCGGAGAA 191

Qy        106 TrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSerSer 125
                 ||||||||||:::|||   |||||||||      |||||||||   ::::::
Db        190 AGCGTCGCGATTCTCTCCGGCGGCGCGCTCCTAGCTCTGGCCTTCGAGCATTTGACGGAA 131

Qy        126 AspAspArgLeuThrAspGluValLysIleLysValLeuGlnArgLeuSerIleAlaSer 145
                 |||   ::::::   |||      ::::::   ::::::   |||:::   :::
Db        130 GCTGAC---GTGTCGCCTGAGAGAATGGTTAGAGCGGTTAAAGAACTGGCGAAGTCTATA 74

Qy        146 GlyHisValGlyMetValGlyGlyGlnMetLeuAspMetGlnSerGluGly 162
                 |||      |||:::|||   ||||||   :::|||:::   |||||||||
Db         73 GGGACGAAAGGGCTCGTGGCGGGACAAGCGATGGATTTGAGCAGCGAAGGT 23

RESULT 15
BI717621
LOCUS       BI717621                 666 bp    mRNA    linear   EST 19-SEP-2001
DEFINITION  1031021A08.y1 C. reinhardtii CC-1690, Stress II (normalized),
            Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION   BI717621
VERSION     BI717621.1  GI:15693316
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii.
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 666)
  AUTHORS   Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Lefebvre
            ,P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
  TITLE     Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants. Project: 1031
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Charles Hauser
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu.
FEATURES             Location/Qualifiers
     source          1. .666
                     /organism="Chlamydomonas reinhardtii"
                     /strain="CC-1690 wild type mt+ 21gr"
                     /db_xref="taxon:3055"
                     /clone_lib="C. reinhardtii CC-1690, Stress II (normalized
                     ), Lambda Zap II"
                     /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
                     XhoI; Stress condition II library, constructed by John
                     Davies and Jeffrey McDermott, combines cDNAs from CC-1690
                     cells grown to mid-log phase in TAP (NH4+ - containing)
                     and shifted to TAP - NO3- (24hrs); H2 production
                     conditions (0, 12hr, 24hr) see Melis et al.,(2000) Plant
                     Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
                     sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
                     PolyA mRNA was purified from each sample, pooled and cDNA
                     synthesized. The cDNA was directionally cloned into lambda
                     Zap II (Stratagene) in the EcoRI (5') and XhoRI (3')
                     sites. pBluescript II SK- plasmids were excised from the
                     lambda ZAP clones by superinfection with ExAssist
                     (Stratagene) phage. The library was normalized using
                     method 4 described in Bonaldo et al., (1996) Genome
                     Research 6: 791-806."
BASE COUNT     126 a    211 c    197 g    132 t
ORIGIN

Alignment Scores:
Pred. No.:              1.05e-27      Length:        666
```

Score:	299.00	Matches:	64
Percent Similarity:	68.91%	Conservative:	18
Best Local Similarity:	53.78%	Mismatches:	29
Query Match:	20.58%	Indels:	8
DB:	13	Gaps:	2

US-09-925-637-64 (1-287) x BI717621 (1-666)

```
Qy          6 MetAsnLysLeuIleAspGluValAsnAsnGluLeuSerValAlaIleAsnLysSerVal 25
              :::|||||||   :::|||||||                        |||:::
Db        331 GTGAACAAGGCCATGGACGAG--------------------GCCGTGCCGCTGAAGTAC 369

Qy         26 MetAspThrGlnLeuGluGluSerMetLeuTyrSerLeuAsnAlaGlyGlyLysArgIle 45
              :::|||   |||   |||||||||   |||||||||   |||||||||||||||:::
Db        370 CCGGAGACT---CTGAACGAGAGCATGCGCTACTCACTGCTGGCTGGCGGCAAGCGCGTG 426

Qy         46 ArgProValLeuLeuLeuLeuThrLeuAspSerLeuAsnThrGluTyrGluLeuGlyMet 65
              ||||||   |||   |||         :::   :::      :::          :::
Db        427 CGCCCAGCTCTCTGCCTGGCGGCCTGCGAGCTGGTTGGCGGCGACATCCACGCCGCGCTG 486

Qy         66 LysSerAlaIleAlaLeuGluMetIleHisThrTyrSerLeuIleHisAspAspLeuPro 85
              :::|||   |||:::||||||:::||||||   ||||||||||||||||||||||||
Db        487 CCTGCTGCGTGCGCCATGGAGATGGTGCACACCATGTCGCTCATCCATGACGACCTGCCC 546

Qy         86 AlaMetAspAsnAspAspTyrArgArgGlyLysLeuThrAsnHisLysValTyrGlyGlu 105
              :::|||||||||||||||:::|||||||||:::   ||||||||||||||||||||||
Db        547 TCCATGGACAACGACGACTTCCGCCGCGGCCGCCCCACCAACCACAAGGTGTACGGCGAG 606

Qy        106 TrpThrAlaIleLeuAlaGlyAspAlaLeuLeuThrLysAlaPheGluLeuIleSer 124
                 |||||||||||||||||||||||||:::   :::||||||   |||:::
Db        607 GACATTGCCATCCTGGCGGGCGACGCGCTGCTGTCGTTCTCGTTTGAGCACATTGCG 663
```

Search completed: May 30, 2003, 21:50:33
Job time : 1562 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 30, 2003, 17:35:15 ; Search time 2483 Seconds
(without alignments)
10091.629 Million cell updates/sec

Title: US-09-925-637-63
Perfect score: 861
Sequence: 1 atgacgaatctaccgatgaa.....tattagaaatcgttgattta 861

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

d size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	810	94.1	333750	1	AP004827	AP004827 Staphyloc
c 2	657	76.3	301550	1	AP003134	AP003134 Staphyloc
c 3	657	76.3	346900	1	AP003362	AP003362 Staphyloc
4	606	70.4	882	6	AR106466	AR106466 Sequence
5	25	2.9	783	6	AX141501	AX141501 Sequence
6	25	2.9	3444	1	AF270104	AF270104 Staphyloc
7	25	2.9	3444	6	AX145422	AX145422 Sequence
c 8	25	2.9	3454	1	AF269958	AF269958 Staphyloc
c 9	25	2.9	3454	6	AX145276	AX145276 Sequence
c 10	25	2.9	4045	1	AF269889	AF269889 Staphyloc
c 11	25	2.9	4045	6	AX145207	AX145207 Sequence
12	25	2.9	111507	9	HS238J17	Z98753 Human DNA s
13	22	2.6	227165	2	AC096106	AC096106 Rattus no
c 14	21	2.4	135599	8	CPU30821	U30821 Cyanophora
15	21	2.4	157270	9	AF236874	AF236874 Homo sapi
16	21	2.4	158344	10	AL133401	AL133401 Mouse DNA
17	21	2.4	159958	9	AP000475	AP000475 Homo sapi
18	21	2.4	170630	2	AC034163	AC034163 Homo sapi
19	21	2.4	171252	2	AC129764	AC129764 Rattus no
20	21	2.4	172663	9	AF241725	AF241725 Homo sapi
c 21	21	2.4	180650	9	AC109474	AC109474 Homo sapi
c 22	21	2.4	200029	2	AC101844	AC101844 Mus muscu
23	21	2.4	340000	9	HS21C027	AL163227 Homo sapi
24	20	2.3	508	10	AF352782	AF352782 Mus muscu
25	20	2.3	884	10	MUSSPKC	J04596 Mouse plate
c 26	20	2.3	3008	5	AF254639	AF254639 Danio rer
27	20	2.3	3282	1	STYPBCDEF	L31538 Salmonella
28	20	2.3	4942	10	S48768	S48768 SP-A=pulmon
29	20	2.3	10029	1	AE006320	AE006320 Lactococc
c 30	20	2.3	22779	1	AE008791	AE008791 Salmonell
c 31	20	2.3	29303	3	CEF57G4	Z81554 Caenorhabdi
c 32	20	2.3	32612	3	CET05D4	Z81115 Caenorhabdi
33	20	2.3	70446	9	AC107299	AC107299 Homo sapi
34	20	2.3	103308	9	AC092491	AC092491 Homo sapi
c 35	20	2.3	105208	10	AL671897	AL671897 Mouse DNA
36	20	2.3	118135	9	AC117477	AC117477 Homo sapi
37	20	2.3	126117	2	AL157949	AL157949 Homo sapi
c 38	20	2.3	126954	9	AL365440	AL365440 Human DNA
c 39	20	2.3	129584	2	AC101911	AC101911 Mus muscu
c 40	20	2.3	140765	9	AC021301	AC021301 Homo sapi
41	20	2.3	146058	9	AC003668	AC003668 Homo sapi
c 42	20	2.3	149777	2	AP003573	AP003573 Oryza sat
43	20	2.3	153369	9	AL592047	AL592047 Human DNA
c 44	20	2.3	156305	2	AC120452	AC120452 Rattus no
c 45	20	2.3	158151	9	CNS05TDT	AL357094 Human chr

ALIGNMENTS

RESULT 1
AP004827/c
LOCUS AP004827 333750 bp DNA linear BCT 02-JUL-2002
DEFINITION Staphylococcus aureus subsp. aureus MW2 DNA, complete genome, strain:MW2, section 6/10.
ACCESSION AP004827 BA000033
VERSION AP004827.1 GI:21204509
KEYWORDS .
SOURCE Staphylococcus aureus subsp. aureus MW2 (strain:MW2) DNA.
ORGANISM Staphylococcus aureus subsp. aureus MW2
Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE 1
AUTHORS Baba,T., Takeuchi,F., Kuroda,M., Yuzawa,H., Aoki,K., Oguchi,A., Nagai,Y., Iwama,N., Asano,K., Naimi,T., Kuroda,H., Cui,L., Yamamoto,K. and Hiramatsu,K.

```
  TITLE     Genome and virulence determinants of high virulence
            community-acquired MRSA
  JOURNAL   Lancet 359 (9320), 1819-1827 (2002)
  MEDLINE   22040717
   PUBMED   12044378
REFERENCE   2  (bases 1 to 333750)
  AUTHORS   Aoki,K., Oguchi,A., Nagai,Y., Asano,K., Iwama,N., Baba,T.,
            Kuroda,M., Hiramatsu,K. and Kikuchi,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAR-2002) Akio Oguchi, National Institute of
            Technology and Evaluation, Biotechnology Center; 2Chome 49-10
            Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
            (E-mail:oguchi@nite.go.jp, URL:http://www.bio.nite.go.jp/,
            Tel:81-3-3481-8423, Fax:81-3-3481-8424)
FEATURES             Location/Qualifiers
     source          1. .333750
                     /organism="Staphylococcus aureus subsp. aureus MW2"
                     /strain="MW2"
                     /db_xref="taxon:196620"
     gene            complement(278. .619)
                     /gene="MW1341"
     CDS             complement(278. .619)
                     /gene="MW1341"
                     /note="ORFID:MW1341"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAB95206.1"
                     /db_xref="GI:21204510"
                     /translation="MTNITKEVFDNLEQEIDVFAKNKTLGSSEAKPYLDEYHSKIIDY
                     FKQVNDITGNIDFDKLNQYPVVPMNFKERYDYMIERKYHFMGYRQMKTFKTELIKMNA
                     SYQTRLKNKQV"
     gene            complement(624. .1283)
                     /gene="nth"
     CDS             complement(624. .1283)
                     /gene="nth"
                     /note="ORFID:MW1342"
                     /codon_start=1
                     /transl_table=11
                     /product="endonuclease-like protein"
                     /protein_id="BAB95207.1"
                     /db_xref="GI:21204511"
                     /translation="MVSKKKALEMIDVIANMFPDAECELKHDNPFELTIAVLLSAQCT
                     DVLVNRVTTELFKKYKTPEDYLAVSDEELMNDIRSIGLYRNKAKNIKKLCQSLIDQFN
                     GEIPQTHKELESLAGVGRKTANVVMSVAFDEPSLAVDTHVERVSKRLGINRWKDNVRQ
                     VEDRLCSVIPRDRWNRSHHQLIFFGRYHCLARKPKCDICPLLEDCREGQKRYKASLKE
                     A"
     gene            complement(1273. .1959)
                     /gene="MW1343"
     CDS             complement(1273. .1959)
                     /gene="MW1343"
                     /note="ORFID:MW1343
                     hypothetical protein, similar to chromosome replication
                     initiation protein dnaD"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB95208.1"
                     /db_xref="GI:21204512"
                     /translation="MDKYQLKARPVVIRRELLDHYSDLGLDEQDLVILLKLIYASETS
                     NKQPSIELLQKGSTMQPRDITMVIQNLIQRELLELQVQKDEEGRFTEYMNLDPFFEKL
                     SHILKQQSMETKEQNSKEKFKQLFRVLEDTFARPLSPYEIETLNQWIDVDKHDSAIIQ
                     AALDEANSLNKLSFKYMDRILLNWKKNNVKTIDDSRKIREKFNKPKMTHTVKTVPKFD
                     WLNGENLDGK"
     gene            complement(2287. .3579)
                     /gene="asnS"
     CDS             complement(2287. .3579)
                     /gene="asnS"
                     /note="ORFID:MW1344"
                     /codon_start=1
                     /transl_table=11
                     /product="asparaginyl-tRNA synthetase"
                     /protein_id="BAB95209.1"
                     /db_xref="GI:21204513"
                     /translation="MKTTIKQAKDHLNQDVTIGAWLTNKRSSGKIAFLQLRDGTGFMQ
                     GVVVKSEVDEEVFKLAKEITQESSLYVTGTITEDNRSDLGYEMQVKSIEVISEAHDYP
                     ITPKNHGTEFLMDHRHLWLRSKKQHAVMKIRNEVIRATYEFFNKDGFTKVDPPILTAS
                     APEGTSELFHTKYFDQDAFLSQSGQLYLEAAAMAHGKVFSFGPTFRAEKSKTRRHLIE
                     FWMIEGEMAFTNHAESLEIQEQYVTHVVKSVLENCKLELKILERDTSKLEKVATPFPR
                     ISYDDAIEFLKSEGFDDIEWGEDFGAPHETAIANHYDLPVFITNYPTKIKPFYMQPNP
                     ENEETVLCADLIAPEGYGEIIGGSERVDDLELLEQRVKEHGLDEEAYSYYLDLRRYGS
                     VPHCGFGLGLERTVAWISGVEHVRETAPFPRLLNRLYP"
     gene            complement(3901. .6594)
                     /gene="dinG"
     CDS             complement(3901. .6594)
                     /gene="dinG"
                     /note="ORFID:MW1345"
                     /codon_start=1
                     /transl_table=11
                     /product="probable ATP-dependent DNA helicase dinG"
                     /protein_id="BAB95210.1"
                     /db_xref="GI:21204514"
                     /translation="MGMATYAVVDLETTGNQLDFDDIIQIGITFVRNNQIIDTYHSMI
                     RTNLEIPPFIQALTSIEENMLQQAPYFNQVAQEIYDKIKDCIFVAHNVDFDLNFIKKA
                     FKDCNIQYRPKKVIDTLEIFKIAFPTDKSYQLSELAEAHGITLANAHRADEDAATTAK
                     LMILAFEKFEKLPLDTLKQLYYLSKQLKYDLYDIFFEMVRQYDAKPLDKSYEKFEQII
                     YRKQVDFKKPTTNYNGSLKSLYSKAVDQLGLTYRPQQLYLAETILDQLMHSEKAMIEA
                     SLGSGKSLAYLLAALMYNIETGKHVMISTNTKLLQSQLLEKDIPAMNEALNFKINALL
                     IKSKSDYISLGLISQILKDDTSNYEVNILKMQLLIWITETPSGDIQELNLKGGQKMYF
                     DQKIETYVPARHDVHYYNFIKRNAQNIQIGITNHAHLIHSDVENSIYQLFDDCIVDEA
                     HRLPDYALNQVTNELSYADIKYQLGLIGKNENEKLLKAIDQLEKQRILEKLDIAPIDI
                     FGLKASMNEIHELNEQLFSTIFTIINDSDVYDDDIHRFHNVFTFETKDILKDLHAIID
                     KLNKTLEIFNGISHKTVKSLRKQLLYLKDKFKNIEQSLKAGHTSFISIKNLSQKSTIR
                     LYVKDYAVKDVLTKQVLEKFKSLIFISGTLKFNHSFEAFKQLFNKDVHFNTFEVNTSL
                     QSAKNTSVFIPSDVASYQYKNIDEYVASIVSYIIEYTTITSSKCLVLFTSYKMMHMVQ
                     DMLNELPEFEDYVVLTQQQNQNYKIVQQFNNFDKAILLGTSTFFEGFDFQANGIKCVM
                     IAKLPFMNKHNAKYWLMDSEFTSTFKEYVLPDAVTRFRQGLGRLIRSENDRGIIVSFD
                     DRLINSNYKNFFEQTLENYRQKKGDIQQFGKLLRQIQKKKK"
     gene            complement(6618. .7589)
                     /gene="MW1346"
     CDS             complement(6618. .7589)
                     /gene="MW1346"
                     /note="ORFID:MW1346"
                     /codon_start=1
                     /transl_table=11
                     /product="truncated hypothetical protein"
                     /protein_id="BAB95211.1"
                     /db_xref="GI:21204515"
                     /translation="MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLK
                     LEGCKIDSVNHKGHLLQQLPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKKSL
                     VGNQSSFFILSDEQTKGRGRFNRHWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALG
                     IRDAIQHFSQDEVKVKWPNDIYIDNGKVCGFLTEMVANNDGIEAIICGIGINLTQQLE
                     NFDESIRHRATSVQLHDKNKLDRYQFLERLLQEIEKRYNQFLTLPFSEIREEYIAASN
                     IWNRTLLFTENDKQFKGQAIDLDYDGYLIVRDEAGESHRLISADIDF"
     gene            complement(7576. .8778)
                     /gene="MW1347"
     CDS             complement(7576. .8778)
                     /gene="MW1347"
                     /note="ORFID:MW1347
                     hypothetical protein, similar to poly(A) polymerase"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB95212.1"
                     /db_xref="GI:21204516"
                     /translation="MDKSLFEQARPILEQIQDNGFEAYYVGGSVRDYVMGRNIHDIDI
                     TTSATPDEIESIFSHTIPVGKEHGTINVVFNDENYEVTTFRAEEDYVDHRRPSGVTFV
                     RDLYEDLQRRDFTMNAIAMDTAYKLYDYFDGQQDINNRIIRTVGIAEERFQEDALRMI
                     RCLRFQSQLSFDIATETFEAMRIQMADIKFLSIERIVIELTKLMRGINVEKSFNHLKS
                     LKAFNYMPYFEHLDMNQINVTEAIDLELLIAIVSVKFDINYSLKPLKLSNRQVKDINQ
                     YIQIMNALPSIITKEQLKMFVYDYDTHLIKNVMVAADVIKANDIQGHEPLIVNLQTID
                     ETLHRLPMHNRKDMMVNGGVLMAHLNAKSGPWLKDVLRQIEIAIVTGKVSNEETEILK
                     WVDNHVKI"
     gene            complement(8783. .9925)
                     /gene="MW1348"
     CDS             complement(8783. .9925)
                     /gene="MW1348"
                     /note="ORFID:MW1348
```

```
                     hypothetical protein, similar to lipopolysaccharide
                     biosynthesis-related pr homolog"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB95213.1"
                     /db_xref="GI:21204517"
                     /translation="MKIGITCYPSMGGSGIIATELGIKLAERGHEVHFITSNIPFRIR
                     KPLPNMIFHQVEVNQYAVFQYPPYDITLSTKIAEVIKEYGLDLLHMHYAVPHAICGIL
                     AREMSGKDIKIMTTLHGTDITVLGYDHSLQGAIKFGIEKSDIVTSVSKSLAQETHEII
                     ETNKEIIPIYNFVRENEFPTKHNTALKSQFGIAPDEKVLIHVSNFRQVKRIDTIIETF
                     AKVREKIPSKLILLGDGPELVPMRQLTKELNVEEDVLFLGKQDCVSEFYQLSDLVLLL
                     SEKESFGLTLLEAMKTGVVPIGSNAGGIKEVIKHGETGFVVDVGDCDSASDYAIRLLE
                     DKVLYNKLQKNMLADIAERFGSELITDQYEYYYQKMLNEHNKSKGE"
     gene            complement(10167. .10484)
                     /gene="MW1349"
     CDS             complement(10167. .10484)
                     /gene="MW1349"
                     /note="ORFID:MW1349"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB95214.1"
                     /db_xref="GI:21204518"
                     /translation="MKSMVEMQREVDEYIGQFKTGYFSPLANLARLTEEVGELAREIN
                     HTYGEKKKKDSEEANTIKAELGDNLFVLLCLANSMGIDMTESFNETMEKFNTRDKNRF
                     ERK"
     gene            complement(10820. .11518)
                     /gene="MW1350"
     CDS             complement(10820. .11518)
                     /gene="MW1350"
                     /note="ORFID:MW1350"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB95215.1"
                     /db_xref="GI:21204519"
                     /translation="MSSIFTMIIYFVILMVLPLWAQHKVKSNYEKYSQVRSTSGKTGR
                     EVALEILHANGIYDVEVVKGEGFLTDHYDPKKKVVSLSPANYDRPSVAGTAIAAHEVG
                     HAIQDHQGYWFLRFRAALVPVANLGSSLSYMIIMLGIILTAIGSAFGSTALWIGAGLM
                     SLAVLFSIVTLPVEFDASSRAMKQITALNIVNEKEYKHARKVLSAAAMTYVAATAVAV
                     AELVRIILLARSSD"

Query Match          94.1%;  Score 810;  DB 1;  Length 333750;
Best Local Similarity  99.9%;  Pred. No. 0;
Matches  860;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118374 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 118315

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118314 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 118255

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118254 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 118195

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118194 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 118135

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118134 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 118075

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118074 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 118015

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118014 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 117955

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117954 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 117895

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117894 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 117835

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117834 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAA 117775

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117774 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 117715

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117714 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 117655

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117654 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 117595

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117594 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 117535

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||||||||||||||||||||
Db     117534 TTATTAGAAATCGTTGATTTA 117514


RESULT 2
AP003134/c
LOCUS       AP003134               301550 bp    DNA     linear   BCT 02-JUL-2002
DEFINITION  Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
            genome, section 6/10.
ACCESSION   AP003134 BA000018
VERSION     AP003134.2  GI:14349226
KEYWORDS    .
SOURCE      Staphylococcus aureus subsp. aureus N315 (sub_species:aureus N315,
            strain:N315) DNA.
  ORGANISM  Staphylococcus aureus subsp. aureus N315
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
            Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M.,
            Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
            Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
            Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
            Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
            Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
  TITLE     Whole genome sequencing of meticillin-resistant Staphylococcus
            aureus
  JOURNAL   Lancet 357 (9264), 1225-1240 (2001)
  MEDLINE   21311952
   PUBMED   11418146
REFERENCE   2  (bases 1 to 301550)
  AUTHORS   Aoki,K., Oguchi,A., Hosoyama,A., Nagai,Y., Kuroda,M., Hiramatsu,K.
            and Kikuchi,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JAN-2001) Akio Oguchi, National Institute of
            Technology and Evaluation, Biotechnology Center; 2Chome 49-10
            Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
            (E-mail:oguchi@nite.go.jp, URL:http://www.bio.nite.go.jp/,
            Tel:81-3-3481-8423, Fax:81-3-3481-8424)
COMMENT     On Jun 12, 2001 this sequence version replaced gi:13701258.
FEATURES             Location/Qualifiers
     source          1. .301550
                     /organism="Staphylococcus aureus subsp. aureus N315"
                     /strain="N315"
```

```
                /sub_species="aureus N315"
                /db_xref="taxon:158879"
     gene       complement(249. .947)
                /gene="SA1293"
     CDS        complement(249. .947)
                /gene="SA1293"
                /note="ORFID:SA1293"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="BAB42553.1"
                /db_xref="GI:13701259"
                /translation="MSSIFTMIIYFVILMVLPLWAQHKVKSNYEKYSQVRSTSGKTGR
                EVALEILHANGIYDVEVVKGEGFLTDHYDPKKKVVSLSPANYDRPSVAGTAIAAHEVG
                HAIQDHQGYWFLRFRAALVPVANLGSSLSYMIIMLGIILTAIGSAFGSTALWIGAGLM
                SLAVLFSIVTLPVEFDASSRAMKQITALNIVNEKEYKHARKVLSAAAMTYVAATAVAV
                AELVRIILIARSSD"
     gene       complement(1001. .1588)
                /gene="SA1294"
     CDS        complement(1001. .1588)
                /gene="SA1294"
                /note="ORFID:SA1294"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="BAB42554.1"
                /db_xref="GI:13701260"
                /translation="MTINTFWQYTLYQRSWLMMLLICNILGMIYGYIWYGEQLSHTPW
                HFKIFVPDSPTAILYLVISISLILIQKQNSIIEALAFVTLFKYGIWAVIMNILFIIEQ
                GDITVNGLVLMFSHSIMAVQAIYFYPRFKRSMIGISVAMIWVFLNDYIDYFHLQFPYY
                DFITTHVWQIGVLSCCLSVFGLLLYIELNKLLKCK"
     gene       complement(1578. .2153)
                /gene="SA1295"
     CDS        complement(1578. .2153)
                /gene="SA1295"
                /note="ORFID:SA1295"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="BAB42555.1"
                /db_xref="GI:13701261"
                /translation="MSETLNQIKESFIEYLLFQYRFKSRIAVWVLNYIKVNEAKLANI
                HFVDTKINHHETLEIAEVGSHASAIQFTKRNIKLMNTNEIFDYIANHNCAFDIQIHFA
                NVSKREQRLDDLIVAQLTESPSYQTYLHDLNSMAIDRHKHALLIDYLLHNIDLSLQMN
                EKQRFYQLTQILNTLKLVNKHNQFEDLADDD"
     gene       complement(2167. .3411)
                /gene="SA1296"
     CDS        complement(2167. .3411)
                /gene="SA1296"
                /note="ORFID:SA1296"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="BAB42556.1"
                /db_xref="GI:13701262"
                /translation="MEDIYKLIDDINLQKLENLDSRVNEAITTDNDDALFILGETLYN
                FGLMPQGLEVFRVLYHKYPDESELLIYFIEGLMSENQTDEALEYLSYVEPSPEKLMLE
                ADLYQQINMMEVAIDKLQEALELEPNDPIIHFALAEMLYYDGQYLRATSEYETVLETG
                EYQVNGVNLFSRMADCSLQSGNYSDAIRLYDEINEDEMTSEDYLKKAISYDKNDITQE
                AIKIMTTLLSKDPDYIQGYLYLQSLYENEKNYPDAIETGKEGLRLSQFYKELMYTTGC
                LEIEHGDANEGVLLLKQALEVDNAYQEPLLILSDLYRNEEDYEAIIELLTYVDEEDLD
                PTFMWHLAFAYGQEERDKEAQHFFELAYPTMKTNIDFMSDYYFYLIEIGQKEQAITIL
                NQLLELEPSNENWHDELSRLQS"
     gene       complement(3418. .4716)
                /gene="aroA"
     CDS        complement(3418. .4716)
                /gene="aroA"
                /note="ORFID:SA1297"
                /codon_start=1
                /transl_table=11
                /product="3-phosphoshikimate 1-carboxyvinyltransferase"
                /protein_id="BAB42557.1"
                /db_xref="GI:13701263"
                /translation="MVSEQIIDISGPLKGEIEVPGDKSMTHRAIMLASLAEGTSNIYK
                PLLGEDCRRTMDIFRLLGVDIKEDEDKLVVNSPGYKAFKTPHQVLYTGNSGTTTRLLA
                GLLSGLGIESVLSGDVSIGKRPMDRVLRPLKLMDANIEGIEDNYTPLIIKPSVIKGIN
                YQMEVASAQVKSAILFASLFSNDTTVIKELDVSRNHTETMFRHFNIPIEAERLSITTT
                PDAIQHIKPADFHVPGDISSAAFFIVAALITPESDVTIHNVGINPTRSGIIDIVEKMG
                GNIQLFNQTTGAEPTASIRIQYTPMLQPITIEGELVPKAIDELPVIALLCTQAVGTST
                IKDAEELKVKETNRIDTTADMLNLLGFELQPTNDGLIIHPSEFKTNATVDSLTDHRIG
                MMLAVASLLSSEPVKIKQFDAVNVSFPGFLPKLKLLENEG"
     gene       complement(4726. .5790)
                /gene="aroB"
     CDS        complement(4726. .5790)
                /gene="aroB"
                /note="ORFID:SA1298"
                /codon_start=1
                /transl_table=11
                /product="3-dehydroquinate synthase"
                /protein_id="BAB42558.1"
                /db_xref="GI:13701264"
                /translation="MKLQTTYPSNNYPIFVEHGAIDHISTYIDQFDQSFILIDEHVNQ
                YFADKFDDILSYENVHKVIIPAGEKTKTFEQYQETLEYILSHHVTRNTAIIAVGGGAT
                GDFAGFVAATLLRGVHFIQVPTTILAHDSSVGGKVGINSKQGKNLIGAFYRPTAVIYD
                LDFLKTLPFEQILSGYAEVYKHALLNGESTTQEIEQHFKDREILQSLNGMDKYIAKGI
                ETKLDIVVADEKEQGVRKFLNLGHTFGHAVEYNHKIAHGHAVMIGIIYQFIVANILFN
                SNHDIQHYINYLTKLGYPLETITDIDFETIYQYMLSDKKNDKQGVQMVLIKHFGDIVV
                QHIDQTTLQHACEQLKTYFK"
     gene       complement(5816. .6982)
                /gene="aroC"
     CDS        complement(5816. .6982)
                /gene="aroC"
                /note="ORFID:SA1299"
                /codon_start=1
                /transl_table=11
                /product="chorismate synthase"
                /protein_id="BAB42559.1"
                /db_xref="GI:13701265"
                /translation="MRYLTSGESHGPQLTVIVEGIPANLEIKVEDINKEMEKRQGGYG
                RGRRMQIEKDTVEIVSGVRNGYTLGSPITMVVTNDDFTHWRKIMGAAPISEEERENMK
                RTITKPRPGHADLVGGMKYNHRDLRNVLERSSARETAARVAVGALCKVLLQQLDIDIY
                SRVVEIGGIKDKDFYDSETFKANLDRNDVRVIDDSIAQAMRDKIDEAKNEGDSIGGVV
                QVVVENMPVGVGSYVHYDRKLDGKIAQGVVSINAFKGVSFGEGFKAAEKPGSEIQDEI
                LYNSEIGYYRGSNHLGGLEGGMSNGMPIIVNGVMKPIPTLYKPLNSVDINTKEDFKAT
                IERSDSCAVPAASIVCEHVVAFEIAKALLEEFQSNHIEQLQQQIADRRQLNVEF"
     gene       complement(7368. .7568)
                /gene="SA1300"
     CDS        complement(7368. .7568)
                /gene="SA1300"
                /note="ORFID:SA1300"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB42560.1"
                /db_xref="GI:13701266"
                /translation="MKNIISIILGILMFLKLMELLYGAIFLDKPLNPITKIIFILTLI
                YIFYVLVKELIIFLKSKYNKSA"
     gene       complement(7597. .7722)
                /gene="SAS045"
     CDS        complement(7597. .7722)
                /gene="SAS045"
                /note="ORFID:SAS045"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB42561.1"
                /db_xref="GI:13701267"
                /translation="MWLVIMIIHNKARFLCCSKLLEKFYSFLVTRIKKRGIFFIK"
     gene       complement(7777. .8226)
                /gene="ndk"
     CDS        complement(7777. .8226)
                /gene="ndk"
                /note="ORFID:SA1301"
                /codon_start=1
                /transl_table=11
                /product="nucleoside diphosphate kinase"
```

```
                     /protein_id="BAB42562.1"
                     /db_xref="GI:13701268"
                     /translation="MERTFLMIKPDAVQRNLIGEVISRIERKGLKLVGGKLMQVPMEL
                     AETHYGEHQGKPFYNDLISFITSAPVFAMVVEGEDAVNVSRHIIGSTNPSEASPGSIR
                     GDLGLTVGRNIIHGSDSLESAEREINLWFNENEITSYASPRDAWLYE"
     gene            complement(8318. .9277)
                     /gene="gerCC"
     CDS             complement(8318. .9277)
                     /gene="gerCC"
                     /note="ORFID:SA1302"
                     /codon_start=1
                     /transl_table=11
                     /product="heptaprenyl diphosphate syntase component II"
                     /protein_id="BAB42563.1"
                     /db_xref="GI:13701269"
                     /translation="MAKLNMNNEIKKVEQRLEKAIKSKDSVLEQASLHLLSSGGKRVR
                     PAFVILSSQFGKDEQTSEQTYQVAVALELIHMATLVHDDVIDKSDKRRGKLTISKKWD

  Query Match           76.3%;  Score 657;  DB 1;  Length 301550;
  Best Local Similarity  99.5%;  Pred. No. 0;
  Matches  857;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64130 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 64071

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db      64070 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGCATGTTGTATTCATTAAATGCT 64011

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64010 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 63951

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db      63950 TATGAGTTAGGTGTGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 63891

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db      63890 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCAT 63831

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63830 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 63771

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63770 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 63711

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63710 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 63651

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63650 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 63591

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db      63590 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAA 63531

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63530 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 63471

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63470 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 63411

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63410 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 63351

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63350 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 63291

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||||||||||||||||||||
Db      63290 TTATTAGAAATCGTTGATTTA 63270


RESULT 3
AP003362/c
LOCUS       AP003362              346900 bp    DNA     linear   BCT 07-FEB-2002
DEFINITION  Staphylococcus aureus subsp. aureus Mu50 genomic DNA, complete
            sequence, section 5/9.
ACCESSION   AP003362 BA000017
VERSION     AP003362.2  GI:14247083
KEYWORDS    .
SOURCE      Staphylococcus aureus subsp. aureus Mu50 (sub_species:aureus Mu50,
            strain:Mu50) DNA.
  ORGANISM  Staphylococcus aureus subsp. aureus Mu50
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
            Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M.,
            Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
            Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
            Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
            Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
            Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
  TITLE     Whole genome sequencing of meticillin-resistant Staphylococcus
            aureus
  JOURNAL   Lancet 357 (9264), 1225-1240 (2001)
  MEDLINE   21311952
REFERENCE   2  (bases 1 to 346900)
  AUTHORS   Ohta,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Toshiko Ohta, University of Tsukuba College
            of Medical Technology and Nursing, Department of Medical
            Technology; 1-1-1 Ten-noodai, Tsukuba, Ibaraki 305-8577, Japan
            (E-mail:tohta@sakura.cc.tsukuba.ac.jp, Tel:81-298-53-3454,
            Fax:81-298-53-3454)
COMMENT     On May 29, 2001 this sequence version replaced gi:13875626.
FEATURES             Location/Qualifiers
     source          1. .346900
                     /organism="Staphylococcus aureus subsp. aureus Mu50"
                     /strain="Mu50"
                     /sub_species="aureus Mu50"
                     /db_xref="taxon:158878"
     gene            12. .263
                     /gene="SAV1313"
     CDS             12. .263
                     /gene="SAV1313"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAB57475.1"
                     /db_xref="GI:14247084"
                     /translation="MLTLIKLKEDEQVIIYEYIPEDDISNGKGSVTFNKKDAEVIDFS
                     LSEIENEEYFMLYRNKSFSVVRDFIEKQEFPENYKIAWY"
     gene            987. .1181
                     /gene="SAV1314"
     CDS             987. .1181
                     /gene="SAV1314"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAB57476.1"
                     /db_xref="GI:14247085"
                     /translation="MNLQSYIGKLVKLTLTNNKILIGKVIDFDDKVDNFDGYNSIEID
                     TGRITYDISENKIKTIVLQK"
     gene            complement(1725. .2750)
                     /gene="SAV1315"
```

```
CDS             complement(1725. .2750)
                /gene="SAV1315"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="BAB57477.1"
                /db_xref="GI:14247086"
                /translation="MISFENDYLEGAHEKVLKRLVDTNLVQASGYGFDQFTAQAIEKI
                KDTIDCPNATIRFLVGGTQTNQVVINSMLESYEGVISADTGHVAVHEGGAIEYSGHKV
                ITIPSKEGKVSASDVETYMETFKSDFKKDHMVFPGMVYISHPTEYGTLYSKSELEELC
                KVCKQYQLPLFMDGARLGYGLMSDQSDMTIKDIAKYCDVFYIGGTKIGALCGEAIVFT
                KNNEPKQFTTRIKHHGALLAKGRLTGIQFLELFTDNLYFNISRHAIEMANKMKDGFKN
                KGYRLYFDSPTNQQFFILSNEKIAELEQKVKFAVWEKYDDQHRVVRFATSWATTEENL
                NKLLELI"
gene            3020. .3217
                /gene="SAV1316"
CDS             3020. .3217
                /gene="SAV1316"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB57478.1"
                /db_xref="GI:14247087"
                /translation="MEQIKLKTFTAETLELLEKNINAFLSSEEATNLKLVNITIKEIE
                ERTFPNNEEEFNAILTLSVNK"
gene            3238. .4755
                /gene="SAV1317"
CDS             3238. .4755
                /gene="SAV1317"
                /codon_start=1
                /transl_table=11
                /product="cardiolipin synthetase homolog"
                /protein_id="BAB57479.1"
                /db_xref="GI:14247088"
                /translation="MFFVPFFIGEKTMRYTFSNDLGTLFTIILAIGFIINLVLAFIII
                FLERNRRTASSTWAWLFVLFVLPLIGFILYLFFGRTVSARKLNKNNGNVLTDFDGLLK
                QQIESFDKGNYGTDNKQVQKHHDLVSMLLMDQDGFLTENNKVDHFIDGNDLYDQVLKD
                IKNAKEYIHLEYYTFALDGLGKRILHALEEKLKQGLEVKILYDDVGSKNVKMANFDHF
                KSLGGEVEAFFASKLPLLNFRMNNRNHRKIIVIDGQLGYVGGFNIGDEYLGLGKLGYW
                RDTHLRIQGDAVDALQLRFILDWNSQAHRPQFEYDVKYFPKKNGPLGNSPIQIAASGP
                ASDWHQIEYGYTKMIMSAKKSVYLQSPYFIPDNSYINAIKIAAKSGVDVHLMIPCKPD
                HPLVYWATFSNASDLLSSGVKIYTYENGFIHSKMCLIDDEIVSVGTANMDFRSFELNF
                EVNAFVYDENLAKDLRVAYEHDITKSKQLTKESYANRPLSVKFKESLAKLVSPIL"
gene            4939. .5436
                /gene="SAV1318"
CDS             4939. .5436
                /gene="SAV1318"
                /note="ATP-binding protein"
                /codon_start=1
                /transl_table=11
                /product="ABC transporter homolog"
                /protein_id="BAB57480.1"
                /db_xref="GI:14247089"
                /translation="MIQISNINKSFNKRCVLKNISFDIEQGKCIALIGKNGAGKSTLI
                DILIGNVNANSGEIFDKDKLLQSENRSIMFQKTMFPDQLKVIEIINLYQSFYENPLPL
                EEIIELTKFDSSQLNQFVNKLSGGQQRLLDFVLSLIGQPQLILLDEPTSTMDIEIREY
                FWSII"
gene            5461. .5838
                /gene="SAV1319"
CDS             5461. .5838
                /gene="SAV1319"
                /note="ATP-binding protein"
                /codon_start=1
                /transl_table=11
                /product="ABC transporter homolog"
                /protein_id="BAB57481.1"
                /db_xref="GI:14247090"
                /translation="MLYTSHYIEEVERMSDKIILIENGEIILNDSTSHIRTNQQSQIT
                LSDEYIRKLKLDKDDLVIQKNHNGTIKIITSNVNDTILYLQQLHINLDDIEIQKVSIV
                DSYFNNKKQRGSNYDTKLLENRI"
gene            5807. .6538
                /gene="SAV1320"
CDS             5807. .6538
                /gene="SAV1320"
                /note="similar to ABC transporter integral membrane
                protein"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB57482.1"
                /db_xref="GI:14247091"
                /translation="MILSYLKIEFKVIMRKKTTLILSILFPVIFYILFTSILELPEDV
                KPKFYKEYMYSMTVYSLLSFSLLTFPLDIINEKQNEWRQRLMVTPFTFTSYYISKVVK
                TMLQFAIAILVIFMVGHFYKGVAMSAVQWLESGIFLWLGASLLITFGILFSLLNDIQK
                TSALANIVTIGLAVLGGLWFPINTFPNWLQHVAHVLPSYHLRKLGVDIASNHHINLIS
                FAIILLYVLGSIIAVYCISHFKRAE"
gene            6542. .7633
                /gene="SAV1321"
CDS             6542. .7633
                /gene="SAV1321"
                /note="similar to two-component sensor histidine kinase"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB57483.1"
                /db_xref="GI:14247092"
                /translation="MKFLKDTSIAEISSILYLIFPIAGIFFNEVYGPKWLYIISVIVF
                SLSYLILVIVNNRLNTLMFYILLIIHYFIICYFVFSVHPMLSLFFFYSAFAVPFTFKN
                NVKKTATNLFILTMIICTIITYLLYNNYFVAMMVYYVVISLIMLDNFKKMKNREYQKE
                IAEKNRHINTLIAEQERHRIGQDLHDTLGHVFASLSLKSELAYKLIDTDVEKVKAELL
                AINKLSRESLNKVREIIDDVKLPSFIEEIDSIRKVLKDADIDFTFENKELAQVLSPTK
                QSMLVMITREAINNVIKHANASKVHGKLKTVNNHKLLLMIEDDGKGIDSDCEVKSISQ
                RVQHLNGTLAVDSTNGTKIIIEISTGGIA"
gene            7630. .8232
                /gene="SAV1322"
CDS             7630. .8232
                /gene="SAV1322"
                /note="similar to two-component response regulator"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB57484.1"
                /db_xref="GI:14247093"
                /translation="MTSLIIAEDQNMLRQAMVQLIKLHGDFEILADTDNGLDAMKLIE
                EYNPNVVILDIEMPGMTGLEVLAEIRKKHLNIKVIIVTTFKRPGYFEKAVVNDVDAYV
                LKERSIEELVETINKVNNGEKEYSATLMTSFFVDKNPLTPKEQIVLREIGNGLSSKEI
                SEKLFLTDGTVRNYTSVIIDKLFADNRFDAWKKANEKGWI"
gene            complement(8345. .8533)
                /gene="SAV1323"
CDS             complement(8345. .8533)
                /gene="SAV1323"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="BAB57485.1"
                /db_xref="GI:14247094"
                /translation="MNISNENTVTLIAVIIAIIIGIFLQIFFKLPLIVTAVLSILLGI
                FVGFIVYLIVSFYNKRKN"
gene            8672. .9205
                /gene="nuc"
CDS             8672. .9205
                /gene="nuc"
                /note="SAV1324"
                /codon_start=1
                /transl_table=11
                /product="thermonuclease"
                /protein_id="BAB57486.1"
                /db_xref="GI:14247095"
                /translation="MKSNKSLAMIVVAIIIVGVLAFQFMNHTGPFKKGTNHETVQDLN
```

Query Match 76.3%; Score 657; DB 1; Length 346900;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 857; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     251159 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 251100

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     251099 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGCATGTTGTATTCATTAAATGCT 251040

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251039 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 250980

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db     250979 TATGAGTTAGGTGTGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 250920

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     250919 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCAT 250860

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      250859 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 250800

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250799 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 250740

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250739 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 250680

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250679 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 250620

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     250619 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAA 250560

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250559 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 250500

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250499 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 250440

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250439 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 250380

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250379 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 250320

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||||||||||||||||||||
Db     250319 TTATTAGAAATCGTTGATTTA 250299


RESULT 4
AR106466
LOCUS       AR106466                882 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 1 from patent US 6107058.
ACCESSION   AR106466
VERSION     AR106466.1  GI:12820996
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 882)
  AUTHORS   Gwynn,M. and Wilding,E.Imogen.
  TITLE     ispA from Staphylococcus aureus
  JOURNAL   Patent: US 6107058-A 1 22-AUG-2000;

FEATURES             Location/Qualifiers
     source          1. .882
                     /organism="unknown"
BASE COUNT      328 a    118 c    186 g    250 t
ORIGIN

  Query Match           70.4%;  Score 606;  DB 6;  Length 882;
  Best Local Similarity  99.4%;  Pred. No. 0;
  Matches  856;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGCATGTTGTATTCATTAAATGCT 120

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTGTGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              ||||||||||||||||||||||||||||||||||||||||| | ||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAAGGAAATTAACAAATCAT 300

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAA 600

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||||||||||||||||||||
Db        841 TTATTAGAAATCGTTGATTTA 861


RESULT 5
AX141501
```

```
LOCUS       AX141501                783 bp    DNA     linear   PAT 31-MAY-2001
DEFINITION  Sequence 223 from Patent WO0134809.
ACCESSION   AX141501
VERSION     AX141501.1  GI:14281622
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 783)
  AUTHORS   Kimmerly,W.J.
  TITLE     Staphylococcus epidermidis nucleic acids and proteins
  JOURNAL   Patent: WO 0134809-A 223 17-MAY-2001;
            GLAXO GROUP LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .783
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic nucleic acid sequence"
BASE COUNT      297 a     93 c    147 g    246 t
ORIGIN

  Query Match             2.9%;  Score 25;  DB 6;  Length 783;
  Best Local Similarity   100.0%;  Pred. No. 0.068;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        281 GAGGAAAATTAACAAATCATAAAGT 305
              |||||||||||||||||||||||||
Db        182 GAGGAAAATTAACAAATCATAAAGT 206


RESULT 6
AF270104
LOCUS       AF270104                3444 bp    DNA     linear   BCT 01-AUG-2000
DEFINITION  Staphylococcus epidermidis strain SR1 clone step.1049b08 genomic
            sequence.
ACCESSION   AF270104
VERSION     AF270104.1  GI:9624010
KEYWORDS    .
SOURCE      Staphylococcus epidermidis.
  ORGANISM  Staphylococcus epidermidis
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1  (bases 1 to 3444)
  AUTHORS   Kimmerly,W.J., Taylor,J.David., Nelsen,A.J., Godlevski,M.M.,
            Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
            Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,
            Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
            Furdon,P.J.
  TITLE     Transposon-mediated sequencing of the Staphylococcus epidermidis
            genome
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3444)
  AUTHORS   Taylor,J.David., Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M.,
            Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
            Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,
            Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
            Furdon,P.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAY-2000) Departments of Genomic Sciences and
            Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
            Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES             Location/Qualifiers
     source          1. .3444
                     /organism="Staphylococcus epidermidis"
                     /strain="SR1"
                     /db_xref="taxon:1282"
                     /clone="step.1049b08"
BASE COUNT     1365 a    451 c    571 g   1057 t
ORIGIN

  Query Match             2.9%;  Score 25;  DB 1;  Length 3444;
  Best Local Similarity   100.0%;  Pred. No. 0.065;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        281 GAGGAAAATTAACAAATCATAAAGT 305
              |||||||||||||||||||||||||
Db       1046 GAGGAAAATTAACAAATCATAAAGT 1070


RESULT 7
AX145422
LOCUS       AX145422                3444 bp    DNA     linear   PAT 31-MAY-2001
DEFINITION  Sequence 4144 from Patent WO0134809.
ACCESSION   AX145422
VERSION     AX145422.1  GI:14283987
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 3444)
  AUTHORS   Kimmerly,W.J.
  TITLE     Staphylococcus epidermidis nucleic acids and proteins
  JOURNAL   Patent: WO 0134809-A 4144 17-MAY-2001;
            GLAXO GROUP LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .3444
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic nucleic acid sequence"
BASE COUNT     1365 a    451 c    571 g   1057 t
ORIGIN

  Query Match             2.9%;  Score 25;  DB 6;  Length 3444;
  Best Local Similarity   100.0%;  Pred. No. 0.065;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        281 GAGGAAAATTAACAAATCATAAAGT 305
              |||||||||||||||||||||||||
Db       1046 GAGGAAAATTAACAAATCATAAAGT 1070


RESULT 8
AF269958/c
LOCUS       AF269958                3454 bp    DNA     linear   BCT 01-AUG-2000
DEFINITION  Staphylococcus epidermidis strain SR1 clone step.1035g12 genomic
            sequence.
ACCESSION   AF269958
VERSION     AF269958.1  GI:9623860
KEYWORDS    .
SOURCE      Staphylococcus epidermidis.
  ORGANISM  Staphylococcus epidermidis
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1  (bases 1 to 3454)
  AUTHORS   Kimmerly,W.J., Taylor,J.David., Nelsen,A.J., Godlevski,M.M.,
            Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
            Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,
            Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
            Furdon,P.J.
  TITLE     Transposon-mediated sequencing of the Staphylococcus epidermidis
            genome
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3454)
  AUTHORS   Taylor,J.David., Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M.,
            Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
            Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,
            Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
            Furdon,P.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAY-2000) Departments of Genomic Sciences and
            Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
            Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES             Location/Qualifiers
     source          1. .3454
                     /organism="Staphylococcus epidermidis"
                     /strain="SR1"
                     /db_xref="taxon:1282"
                     /clone="step.1035g12"
```

```
BASE COUNT     1005 a    574 c    493 g   1382 t
ORIGIN

  Query Match             2.9%;  Score 25;  DB 1;  Length 3454;
  Best Local Similarity   100.0%;  Pred. No. 0.065;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        281 GAGGAAAATTAACAAATCATAAAGT 305
              |||||||||||||||||||||||||
Db         77 GAGGAAAATTAACAAATCATAAAGT 53


RESULT 9
AX145276/c
LOCUS       AX145276               3454 bp    DNA     linear   PAT 31-MAY-2001
DEFINITION  Sequence 3998 from Patent WO0134809.
ACCESSION   AX145276
VERSION     AX145276.1  GI:14283841
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 3454)
  AUTHORS   Kimmerly,W.J.
  TITLE     Staphylococcus epidermidis nucleic acids and proteins
  JOURNAL   Patent: WO 0134809-A 3998 17-MAY-2001;
            GLAXO GROUP LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .3454
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic nucleic acid sequence"
BASE COUNT     1005 a    574 c    493 g   1382 t
ORIGIN

  Query Match             2.9%;  Score 25;  DB 6;  Length 3454;
  Best Local Similarity   100.0%;  Pred. No. 0.065;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        281 GAGGAAAATTAACAAATCATAAAGT 305
              |||||||||||||||||||||||||
Db         77 GAGGAAAATTAACAAATCATAAAGT 53


RESULT 10
AF269889/c
LOCUS       AF269889               4045 bp    DNA     linear   BCT 01-AUG-2000
DEFINITION  Staphylococcus epidermidis strain SR1 clone step.1028f08 genomic
            sequence.
ACCESSION   AF269889
VERSION     AF269889.1  GI:9623789
KEYWORDS    .
SOURCE      Staphylococcus epidermidis.
  ORGANISM  Staphylococcus epidermidis
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1  (bases 1 to 4045)
  AUTHORS   Kimmerly,W.J., Taylor,J.David., Nelsen,A.J., Godlevski,M.M.,
            Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
            Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,
            Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
            Furdon,P.J.
  TITLE     Transposon-mediated sequencing of the Staphylococcus epidermidis
            genome
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4045)
  AUTHORS   Taylor,J.David., Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M.,
            Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
            Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,
            Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
            Furdon,P.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAY-2000) Departments of Genomic Sciences and
            Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
            Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES             Location/Qualifiers
     source          1. .4045
                     /organism="Staphylococcus epidermidis"
                     /strain="SR1"
                     /db_xref="taxon:1282"
                     /clone="step.1028f08"
BASE COUNT     1249 a    691 c    536 g   1569 t
ORIGIN

  Query Match             2.9%;  Score 25;  DB 1;  Length 4045;
  Best Local Similarity   100.0%;  Pred. No. 0.065;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        281 GAGGAAAATTAACAAATCATAAAGT 305
              |||||||||||||||||||||||||
Db       2033 GAGGAAAATTAACAAATCATAAAGT 2009


RESULT 11
AX145207/c
LOCUS       AX145207               4045 bp    DNA     linear   PAT 31-MAY-2001
DEFINITION  Sequence 3929 from Patent WO0134809.
ACCESSION   AX145207
VERSION     AX145207.1  GI:14283772
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4045)
  AUTHORS   Kimmerly,W.J.
  TITLE     Staphylococcus epidermidis nucleic acids and proteins
  JOURNAL   Patent: WO 0134809-A 3929 17-MAY-2001;
            GLAXO GROUP LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .4045
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic nucleic acid sequence"
BASE COUNT     1249 a    691 c    536 g   1569 t
ORIGIN

  Query Match             2.9%;  Score 25;  DB 6;  Length 4045;
  Best Local Similarity   100.0%;  Pred. No. 0.065;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        281 GAGGAAAATTAACAAATCATAAAGT 305
              |||||||||||||||||||||||||
Db       2033 GAGGAAAATTAACAAATCATAAAGT 2009


RESULT 12
HS238J17
LOCUS       HS238J17               111507 bp    DNA     linear   PRI 23-NOV-1999
DEFINITION  Human DNA sequence from PAC 238J17 on chromosome 6q22. Contains EST
            and STS.
ACCESSION   Z98753
VERSION     Z98753.1  GI:2760552
KEYWORDS    6q22.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 111507)
  AUTHORS   Phillips,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-DEC-1997) Chromosome 6 Project Group
            (http://wwwsanger.ac.uk/HGP/Chr6/) Sanger Centre, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquires:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jan 9, 1998 this sequence version replaced gi:2578075.
            IMPORTANT: This sequence is the entire insert of clone 238J17.
            During sequence assembly data is compared from overlapping clones.
```

Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variations annotated may not be found in the sequence submission
corresponding to the overlapping clone as we submit sequences with
only a small overlap as described above.
This sequence was generated from part of bacterial clone contigs of
human chromosome 6, constructed by the Sanger Centre chromosome 6
mapping group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6/
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The true left end of clone 238J17 is at 1 in this sequence. The
true right end of clone 238J17 is at 111507.
238J17 is from the library RPCI1 constructed at the Roswell Park
Cancer Institute by the group of Pieter de Jong.
For further details see http://bacpac.med.buffalo.edu/.

TURES Location/Qualifiers
source 1. .111507
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6q22"
/clone="RP1-238J17"
/clone_lib="RPCI-1"
repeat_region 2. .816
/note="L1PA16 repeat: matches 816. .1 of consensus"
repeat_region 691. .2583
/note="L1 repeat: matches 5371. .3469 of consensus"
repeat_region 3298. .4092
/note="L1PA15 repeat: matches 904. .108 of consensus"
repeat_region 4093. .4354
/note="AluSg repeat: matches 297. .36 of consensus;
incomplete repeat"
repeat_region 4357. .4880
/note="L1 repeat: matches 5345. .4816 of consensus"
repeat_region 7245. .8436
/note="TIGGER1 repeat: matches 2418. .1166 of consensus"
repeat_region 8462. .8589
/note="AluJb repeat: matches 131. .1 of consensus;
incomplete repeat"
repeat_region 8616. .9664
/note="TIGGER1 repeat: matches 1156. .91 of consensus"
repeat_region 11266. .11319
/note="27 copies of 2 mer 96 % conserved"
repeat_region 11811. .11942
/note="MER45 repeat: matches 178. .59 of consensus"
repeat_region 12679. .12714
/note="18 copies of 2 mer 97 % conserved"
repeat_region 14434. .14578
/note="MLT1G repeat: matches 165. .31 of consensus"
repeat_region 14795. .14930
/note="MLT2FB repeat: matches 1. .140 of consensus"
repeat_region 14941. .15016
/note="MLT2E repeat: matches 282. .357 of consensus"
repeat_region 15061. .15108
/note="MLT2E repeat: matches 350. .397 of consensus"
repeat_region 17447. .18339
/note="L1PA10 repeat: matches 905. .1 of consensus"
repeat_region 18195. .23091
/note="L1 repeat: matches 5390. .593 of consensus"
repeat_region 23607. .23839
/note="L1 repeat: matches 256. .11 of consensus"
repeat_region 27516. .27656
/note="MIR repeat: matches 122. .257 of consensus"
repeat_region 28584. .29013
/note="MLT1C repeat: matches 466. .8 of consensus"
repeat_region 33266. .33895
/note="L1 repeat: matches 4575. .5220 of consensus"
repeat_region 33966. .34245
/note="AluSc repeat: matches 1. .299 of consensus"
repeat_region 34257. .35032
/note="L1MA10 repeat: matches 27. .737 of consensus"
repeat_region 34553. .35348
/note="L1MC1 repeat: matches 255. .1078 of consensus"
repeat_region 36418. .36439
/note="11 copies of 2 mer 100 % conserved"
repeat_region 36461. .36816
/note="THE1B repeat: matches 364. .1 of consensus"
repeat_region 36817. .36912
/note="24 copies of 4 mer 82 % conserved"
repeat_region 36960. .37879
/note="L1 repeat: matches 4471. .5390 of consensus"
repeat_region 37730. .38265
/note="L1PB1 repeat: matches 1. .538 of consensus"
repeat_region 38468. .38491
/note="12 copies of 2 mer 100 % conserved"
repeat_region 38494. .38853
/note="L1PB2 repeat: matches 530. .897 of consensus"
repeat_region 39788. .40054
/note="AluY repeat: matches 38. .301 of consensus;
incomplete repeat"
repeat_region 41359. .41419
/note="AluSq repeat: matches 61. .1 of consensus;
incomplete repeat"
repeat_region 43483. .43627
/note="MIR2 repeat: matches 1. .146 of consensus"
repeat_region 44040. .44431
/note="MLT1B repeat: matches 390. .1 of consensus"
repeat_region 46359. .47262
/note="L1PA2 repeat: matches 893. .1 of consensus"
repeat_region 47119. .51945
/note="L1 repeat: matches 5390. .589 of consensus"
repeat_region 52044. .52239
/note="L1 repeat: matches 566. .381 of consensus"
repeat_region 52464. .52771
/note="L1 repeat: matches 348. .13 of consensus"
repeat_region 58814. .59121
/note="AluJo repeat: matches 1. .299 of consensus"
repeat_region 59600. .59631
/note="16 copies of 2 mer 91 % conserved"
repeat_region 61461. .61626
/note="MIR repeat: matches 261. .91 of consensus"
repeat_region 63140. .63435
/note="AluJo repeat: matches 295. .1 of consensus"
repeat_region 63632. .64009
/note="MSTA repeat: matches 426. .28 of consensus"
repeat_region 64275. .64575
/note="AluJb repeat: matches 300. .1 of consensus"
repeat_region 64703. .65077
/note="THE1C repeat: matches 371. .1 of consensus"
repeat_region 65078. .66224
/note="THE1B-INTERNAL repeat: matches 1580. .437 of
consensus"
repeat_region 66227. .66618
/note="THE1B-INTERNAL repeat: matches 395. .1 of consensus"
repeat_region 66619. .66907
/note="THE1C repeat: matches 371. .87 of consensus"
repeat_region 66922. .67285
/note="THE1B repeat: matches 1. .364 of consensus"
repeat_region 67294. .67357
/note="THE1B repeat: matches 68. .3 of consensus"
repeat_region 69698. .70008
/note="MER2 repeat: matches 3. .345 of consensus"
repeat_region 70394. .70498
/note="L1MA1 repeat: matches 947. .1051 of consensus"
repeat_region 70504. .70798
/note="AluSg repeat: matches 1. .300 of consensus"
repeat_region 70999. .71134
/note="MLT1G repeat: matches 133. .8 of consensus"
repeat_region 71973. .72147
/note="MIR repeat: matches 240. .54 of consensus"
repeat_region 72847. .73211

```
                     /note="THE1B repeat: matches 1. .364 of consensus"
     repeat_region   73212. .74406
                     /note="THE1B-INTERNAL repeat: matches 1. .1157 of
                     consensus"
     repeat_region   74553. .75509
                     /note="MER11A repeat: matches 59. .1090 of consensus"
     repeat_region   75510. .75941
                     /note="THE1B-INTERNAL repeat: matches 1153. .1580 of
                     consensus"
     repeat_region   75942. .76305
                     /note="THE1B repeat: matches 1. .364 of consensus"
     repeat_region   78184. .78607
                     /note="MSTA repeat: matches 424. .1 of consensus"
     repeat_region   80178. .80727
                     /note="MLT2D repeat: matches 545. .1 of consensus"
     repeat_region   82250. .82366
                     /note="MIR repeat: matches 157. .29 of consensus"
     misc_feature    82329. .82669
                     /note="match: STS G05363"
     repeat_region   85010. .85309
                     /note="AluSg repeat: matches 300. .1 of consensus"
     repeat_region   86582. .86876
                     /note="AluY repeat: matches 293. .1 of consensus"
     repeat_region   86901. .87116
                     /note="L1MA3 repeat: matches 614. .399 of consensus"
     repeat_region   86907. .87066
                     /note="L1MD2 repeat: matches 610. .457 of consensus"
     repeat_region   87406. .92840
                     /note="L1 repeat: matches 11. .5390 of consensus"
     repeat_region   92699. .93572
                     /note="L1PA3 repeat: matches 1. .891 of consensus"
     repeat_region   93896. .93921
                     /note="13 copies of 2 mer 100 % conserved"
     repeat_region   94298. .94642

  Query Match             2.9%;  Score 25;  DB 9;  Length 111507;
  Best Local Similarity   100.0%;  Pred. No. 0.059;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        387 AACTGATGAAGTAAAAATAAAAGTT 411
              |||||||||||||||||||||||||
Db      46032 AACTGATGAAGTAAAAATAAAAGTT 46056


RESULT 13
AC096106
LOCUS       AC096106               227165 bp    DNA     linear   HTG 11-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-27L19, *** SEQUENCING IN PROGRESS
            ***, 111 unordered pieces.
ACCESSION   AC096106
VERSION     AC096106.3  GI:21239786
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 227165)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
    TITLE   Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 227165)
  AUTHORS   Worley,K.C.
    TITLE   Direct Submission
  JOURNAL   Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 227165)
  AUTHORS   Worley,K.C.
    TITLE   Direct Submission
  JOURNAL   Submitted (11-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 29, 2002 this sequence version replaced gi:17943790.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GELF
                Center clone name: CH230-27L19
            -------------- Summary Statistics
                Sequencing vector: Plasmid;
                Chemistry: Dye-terminator Big Dye: 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 113322 bases at least Q40
                Consensus quality: 121538 bases at least Q30
                Consensus quality: 128485 bases at least Q20
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 111 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     1469: contig of 1469 bp in length
            *    1470     1569: gap of unknown length
            *    1570     2648: contig of 1079 bp in length
            *    2649     2748: gap of unknown length
            *    2749     4051: contig of 1303 bp in length
            *    4052     4151: gap of unknown length
            *    4152     5750: contig of 1599 bp in length
```

```
*     5751     5850: gap of unknown length
*     5851     6946: contig of 1096 bp in length
*     6947     7046: gap of unknown length
*     7047     8189: contig of 1143 bp in length
*     8190     8289: gap of unknown length
*     8290     9364: contig of 1075 bp in length
*     9365     9464: gap of unknown length
*     9465    10485: contig of 1021 bp in length
*    10486    10585: gap of unknown length
*    10586    12155: contig of 1570 bp in length
*    12156    12255: gap of unknown length
*    12256    13531: contig of 1276 bp in length
*    13532    13631: gap of unknown length
*    13632    15089: contig of 1458 bp in length
*    15090    15189: gap of unknown length
*    15190    16711: contig of 1522 bp in length
*    16712    16811: gap of unknown length
*    16812    17941: contig of 1130 bp in length
*    17942    18041: gap of unknown length
*    18042    19093: contig of 1052 bp in length
*    19094    19193: gap of unknown length
*    19194    20353: contig of 1160 bp in length
*    20354    20453: gap of unknown length
*    20454    21575: contig of 1122 bp in length
*    21576    21675: gap of unknown length
*    21676    23040: contig of 1365 bp in length
*    23041    23140: gap of unknown length
*    23141    24260: contig of 1120 bp in length
*    24261    24360: gap of unknown length
*    24361    25396: contig of 1036 bp in length
*    25397    25496: gap of unknown length
*    25497    27048: contig of 1552 bp in length
*    27049    27148: gap of unknown length
*    27149    28238: contig of 1090 bp in length
*    28239    28338: gap of unknown length
*    28339    30192: contig of 1854 bp in length
*    30193    30292: gap of unknown length
*    30293    31499: contig of 1207 bp in length
*    31500    31599: gap of unknown length
*    31600    33540: contig of 1941 bp in length
*    33541    33640: gap of unknown length
*    33641    35223: contig of 1583 bp in length
*    35224    35323: gap of unknown length
*    35324    36759: contig of 1436 bp in length
*    36760    36859: gap of unknown length
*    36860    38728: contig of 1869 bp in length
*    38729    38828: gap of unknown length
*    38829    40656: contig of 1828 bp in length
*    40657    40756: gap of unknown length
*    40757    42105: contig of 1349 bp in length
*    42106    42205: gap of unknown length
*    42206    43532: contig of 1327 bp in length
*    43533    43632: gap of unknown length
*    43633    45310: contig of 1678 bp in length
*    45311    45410: gap of unknown length
*    45411    46434: contig of 1024 bp in length
*    46435    46534: gap of unknown length
*    46535    47673: contig of 1139 bp in length
*    47674    47773: gap of unknown length
*    47774    49578: contig of 1805 bp in length
*    49579    49678: gap of unknown length
*    49679    51420: contig of 1742 bp in length
*    51421    51520: gap of unknown length
*    51521    52838: contig of 1318 bp in length
*    52839    52938: gap of unknown length
*    52939    54588: contig of 1650 bp in length
*    54589    54688: gap of unknown length
*    54689    55991: contig of 1303 bp in length
*    55992    56091: gap of unknown length
*    56092    57458: contig of 1367 bp in length
*    57459    57558: gap of unknown length
*    57559    58637: contig of 1079 bp in length
*    58638    58737: gap of unknown length
*    58738    60572: contig of 1835 bp in length
*    60573    60672: gap of unknown length
*    60673    61999: contig of 1327 bp in length
*    62000    62099: gap of unknown length
*    62100    63876: contig of 1777 bp in length
*    63877    63976: gap of unknown length
*    63977    65299: contig of 1323 bp in length
*    65300    65399: gap of unknown length
*    65400    66581: contig of 1182 bp in length
*    66582    66681: gap of unknown length
*    66682    67940: contig of 1259 bp in length
*    67941    68040: gap of unknown length
*    68041    69984: contig of 1944 bp in length
*    69985    70084: gap of unknown length
*    70085    71379: contig of 1295 bp in length
*    71380    71479: gap of unknown length
*    71480    73359: contig of 1880 bp in length
*    73360    73459: gap of unknown length
*    73460    74946: contig of 1487 bp in length
*    74947    75046: gap of unknown length
*    75047    76380: contig of 1334 bp in length
*    76381    76480: gap of unknown length
*    76481    78494: contig of 2014 bp in length
*    78495    78594: gap of unknown length
*    78595    79825: contig of 1231 bp in length
*    79826    79925: gap of unknown length
*    79926    81163: contig of 1238 bp in length

  Query Match             2.6%;  Score 22;  DB 2;  Length 227165;
  Best Local Similarity   100.0%;  Pred. No. 2.4;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        387 AACTGATGAAGTAAAAATAAAA 408
              ||||||||||||||||||||||
Db     157861 AACTGATGAAGTAAAAATAAAA 157882


RESULT 14
CPU30821/c
LOCUS       CPU30821               135599 bp    DNA     circular PLN 13-NOV-1995
DEFINITION  Cyanophora paradoxa cyanelle, complete genome.
ACCESSION   U30821
VERSION     U30821.1  GI:1016083
KEYWORDS    .
SOURCE      Cyanophora paradoxa.
  ORGANISM  Cyanelle Cyanophora paradoxa
            Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.
REFERENCE   1  (bases 1 to 135599)
  AUTHORS   Stirewalt,V.L., Michalowski,C.B., Luffelhardt,W., Bohnert,H.J. and
            Bryant,D.A.
  TITLE     Nucleotide sequence of the cyanelle genome from Cyanophora paradoxa
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 135599)
  AUTHORS   Bryant,D.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUL-1995) Donald A. Bryant, Biochemistry and
            Molecular Biology, The Pennsylvania State University, S-234 Frear
            Bldg., University Park, PA 16802, USA
FEATURES             Location/Qualifiers
     source          1. .135599
                     /organism="Cyanophora paradoxa"
                     /organelle="plastid:cyanelle"
                     /strain="Pringsheim LB555"
                     /db_xref="taxon:2762"
     gene            complement(join(134890. .135599,1. .10))
                     /gene="ycf27"
     CDS             complement(join(134890. .135599,1. .10))
                     /gene="ycf27"
                     /function="transcriptional regulator of the OmpR family;
                     target unknown"
                     /codon_start=1
                     /protein_id="AAA81319.1"
                     /db_xref="GI:1016232"
```

```
                /translation="MEINKKKILVVDDEISIRRILETRLSMIGYEVVTAADGEEALTI
                FQLEHPNLVVLDVMMPKLDGYGVCQELRKESDIPIIMLTALGDVADRITGLELGADDY
                VVKPFSPKELEARIRSLLRRTENPTLSTTSTNGENLQIGFLKIDINKRQVFKNGERIR
                LTGMEFSLLELLISKMGEPFSRAQILQEVWGYTPERHIDTRVVDVHISRLRSKLEENP
                SNPDLILTARGIGYLFQNTSN"
repeat_region   1. .11285
                /note="IR-A"
                /rpt_type=inverted
gene            294. .1788
                /gene="rrsA-A"
rRNA            294. .1788
                /gene="rrsA-A"
                /product="16S ribosomal RNA"
gene            1889. .1962
                /gene="trnI-A"
tRNA            1889. .1962
                /gene="trnI-A"
                /product="tRNA-Ile"
                /note="anticodon (GAU)"
gene            1966. .2038
                /gene="trnA-A"
tRNA            1966. .2038
                /gene="trnA-A"
                /product="tRNA-Ala"
                /note="anticodon (UGC)"
gene            2076. .5001
                /gene="rrlA-A"
rRNA            2076. .5001
                /gene="rrlA-A"
                /product="23S ribosomal RNA"
gene            5102. .5219
                /gene="rrfA-A"
rRNA            5102. .5219
                /gene="rrfA-A"
                /product="5S ribosomal RNA"
gene            5727. .7598
                /gene="dnaK-A"
CDS             5727. .7598
                /gene="dnaK-A"
                /function="molecular chaperone; heat-shock protein;
                protein folding"
                /codon_start=1
                /product="DnaK"
                /protein_id="AAA81171.1"
                /db_xref="GI:1016084"
                /translation="MEVKHMGKIVGIDLGTTNSVVAVMEGGKPTVITNAEGFRTTPSV
                VAYTKTGDRLVGQIAKRQAVLNPGNTFYSVKRFIGRKFDEVTEEAKQVPYKVTGTQNV
                RIECPALQTSFTPEEISAQVLRKLAEDASKYIGEPVTQAVITVPAYFNDSQRQATKDA
                GKIAGLEVLRIINEPTAASLAYGLDKKKNETILVFDLGGGTFDVSILEVGDGVFEVLS
                TSGDTHLGGDDFDKKLVDFFIEEFQKLENINLREDQQALQRLTEAAEKAKIELSSVTQ
                TEITLPFLSLTPSGPKHFDTTLNRAKFEELCSTLIDRCRIPVETALKDAKLTKEDIDE
                VVLVGGSTRIPAVRELVQKVLDKVPKQSVNPDEVVAVGAAVQAGVLAGEVKDIVLLDV
                TPLSLGVETLGGVMSKIIVRNTTIPTQKSELFSTATDGQTNVEIHVLQGERGFAKDNK
                SLGTFRLDGIAPAPRGVPQIEVTFDIDANGILSVTAKDKGTGKEKSITITGASTLSQD
                EVERMVQEAEKNAAQDKENRQKIELRNQAESLINQAERQLKEFKLEPNVQAPIEQTIQ
                QLQTALKEDNLEQVQSLCQELQNSLLEIGKQVYSKTETTTPNKNEEDVIDASFSEEK"
gene            7694. .7764
                /gene="trnC-A"
tRNA            7694. .7764
                /gene="trnC-A"
                /product="tRNA-Cys"
                /note="anticodon (GCA)"
gene            complement(8038. .9663)
                /gene="groEL-A"
CDS             complement(8038. .9663)
                /gene="groEL-A"
                /function="molecular chaperone; heat-shock protein;
                protein folding"
                /codon_start=1
                /product="GroEL"
                /protein_id="AAA81172.1"
                /db_xref="GI:1016085"
                /translation="MAKQILYHESARRALEKGMDILAEAVAVTLGPKGRNVVLEKKFG
                APQIVNDGVTIAKEIELEDHIENTGVALIRQAASKTNDVAGDGTTTATVLAYAMVKEG
                MRNVAAGANPMAIKRGIDRATQKVVETIAAHAKPVEDSRAITQVASISAGNDDEVGKM
                IADAIEKVGKEGVISLEEGKSTTTELEITEGMRFDKGYISPYFVTDPERMEVIQKDAM
                FLITDKKITLVQDLVPILEQVARARKPLFIIAEDIEKEALATLVVNKLRGNLNVSAVK
                APGFGDRRKAMMQDIAVLTNTQVISEQTGLRLDSIKLDVLGKARQVNITKDYTTVIAE
                GNEEGVSARCEQIRRQIEETDSAYEKEKLQERLAKLAGGVAVIKVGAATETEMKDKKL
                RLEDAINATKAAVEEGIVPGGGTTLAHLAIELENWVQSNLEHEELIGGMIVARSLTAP
                LKRIVENAGQNGAVIAEQVKDKPFNIGYNAANAQFVDMFEAGIVDPAKVTRSGLQNAA
                SIAGMVLTTECIVVDKAESQQKAASVNRNSFDY"
gene            complement(9727. .10038)
                /gene="groES-A"
CDS             complement(9727. .10038)
                /gene="groES-A"
                /function="molecular chaperone; chaperonin; heat-shock
                protein; protein folding"
                /codon_start=1
                /product="GroES"
                /protein_id="AAA81173.1"
                /db_xref="GI:1016086"
                /translation="MATVTLNVKTVRPLGERVLVKVSQSEEKTAGGILLPDTVKEKPQ
                IGEIIAEGPGRRNDDGSFQPLEVTVNSKVLYSKYAGTDIKLENEEYVLLSEKDILAII
                A"
gene            10419. .11003
                /gene="clpPI"
CDS             10419. .11003
                /gene="clpPI"
                /function="ATP-dependent protease subunit"
                /codon_start=1
                /product="ClpP1 protease subnit"
                /protein_id="AAA81174.1"
                /db_xref="GI:1016087"
                /translation="MTILIQQYTNDENKKLELNNISKLLQNRIILFSQTVDDEICNSI
                VGQLLYLENEDSTKDIRLFINSPGGSVTAGLSVYDTIQNLSVDVSTICFGLAASMGAV
                LLAAGVENKRFAFASSRIMIHQPLSKVEAPWSHLDIQIRNGAYFKNLLNNILSFHTKQ
                ELKQIETDTERDFFLSATEAKQYGIIDHIFINNN"
gene            11515. .11586
                /gene="trnK"
tRNA            11515. .11586
                /gene="trnK"
                /product="tRNA-Lys"
                /note="anticodon (UUU)"
gene            complement(11591. .12109)
                /gene="ycf37"
CDS             complement(11591. .12109)
                /gene="ycf37"
                /codon_start=1
                /protein_id="AAA81175.1"
                /db_xref="GI:1016088"
                /translation="MDNFLFSFYIIFFSLIVLALISFLIQEIQKSRFLDLEITRLRKK
                IYQEKATEKDYNNLASIYIKKKLYSQAQKEFLKILENKNLDKSQLALIYNSLGYICSA
                QEQYELAIEYYKKALFYIPDFILARINLARIFELKNLKTEAINMYQEVLIFDPKNQLA
                KRKLSIIENLIL"
gene            complement(12224. .12457)
                /gene="ycf34"
CDS             complement(12224. .12457)
                /gene="ycf34"
                /codon_start=1
                /protein_id="AAA81176.1"
                /db_xref="GI:1016089"
                /translation="MCICVNCDLLVFCSLYQTIEILHKESLSFELDQIVEFEPRKSNI
                IVNINSSENSIKHEWDVISCESFVLNLSKYNVS"
gene            complement(join(12625. .12674,12907. .12941))
                /gene="trnL-A"
tRNA            complement(join(12625. .12674,12907. .12941))
                /gene="trnL-A"
                /product="tRNA-Leu"
                /note="anticodon (UAA)"
intron          complement(12675. .12906)
                /gene="trnL-A"
                /note="Type-I intron; found in many cyanobacteria and some
                eubacteria"
gene            complement(13099. .13171)
                /gene="trnE"
```

```
     tRNA            complement(13099. .13171)
                     /gene="trnE"
                     /note="anticodon (UUC)"
     gene            13520. .13831
                     /gene="acpP"
     CDS             13520. .13831
                     /gene="acpP"
                     /function="fatty acid biosynthesis"
                     /codon_start=1
                     /product="acyl carrier protein"
                     /protein_id="AAA81177.1"
                     /db_xref="GI:1016090"
                     /translation="MSLSEQAILEQDKNIVSNIVQDILCEQISIEKTELNLDLYIYED
                     LKIDSLDLVEIIKQIEEKFDIKIDDSKILYMNTLEEFIDFTLQTIYMKHGFEYLQNKQ
                     F"
     gene            13893. .14456
                     /gene="ycf4"
     CDS             13893. .14456
                     /gene="ycf4"
                     /function="possible in regulating PS I:PS II ratio"

  Query Match              2.4%;  Score 21;  DB 8;  Length 135599;
  Best Local Similarity   100.0%;  Pred. No. 8.7;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        328 TTAGCAGGTGATGCTTTATTA 348
              |||||||||||||||||||||
Db     129578 TTAGCAGGTGATGCTTTATTA 129558


RESULT 15
AF236874
LOCUS       AF236874               157270 bp    DNA     linear   PRI 02-MAR-2000
DEFINITION  Homo sapiens chromosome 21 map 21q21, BAC B242B02, complete
            sequence.
ACCESSION   AF236874
VERSION     AF236874.1  GI:7141351
KEYWORDS    HTG.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 157270)
  AUTHORS   Blechschmidt,K., Riesselmann,L., Wehrmeyer,S., Baumgart,C.,
            Menzel,U., Dette,M., Jahn,N., Schilhabel,M., Yaspo,M.-L. and
            Rosenthal,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-FEB-2000) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
FEATURES             Location/Qualifiers
     source          1. .157270
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q21"
                     /clone="BAC B242B02"
     repeat_region   complement(1. .134)
                     /rpt_family="MLT1F"
     repeat_region   complement(19. .429)
                     /rpt_family="MLT1E"
     repeat_region   2089. .2400
                     /rpt_family="AluSq"
     repeat_region   2700. .2802
                     /rpt_family="MER81"
     exon            2730. .2840
                     /note="GRAIL"
                     /evidence=not_experimental
     repeat_region   complement(2943. .3223)
                     /rpt_family="MLT1E"
     repeat_region   3798. .4553
                     /rpt_family="L1MCb_5end"
     repeat_region   complement(3824. .4287)
                     /rpt_family="MLT1D"
     exon            4542. .4639
                     /note="GRAIL"
                     /evidence=not_experimental
     repeat_region   complement(4817. .5043)
                     /rpt_family="LOR1a"
     repeat_region   complement(5041. .5106)
                     /rpt_family="LOR1a"
     repeat_region   complement(5112. .5402)
                     /rpt_family="AluSx"
     repeat_region   5648. .6167
                     /rpt_family="L1MCb_5end"
     repeat_region   6180. .6502
                     /rpt_family="L1MEc_5end"
     repeat_region   complement(6613. .7086)
                     /rpt_family="L1PA15"
     repeat_region   complement(7161. .7578)
                     /rpt_family="L1PA14"
     repeat_region   complement(7436. .11435)
                     /rpt_family="L1"
     repeat_region   7861. .7884
                     /note="(GGAA)n"
                     /rpt_family="Simple_repeat"
     exon            complement(8364. .8513)
                     /note="GRAIL"
                     /evidence=not_experimental
     repeat_region   complement(11225. .11846)
                     /rpt_family="L1PA12_5end"
     repeat_region   complement(11516. .12125)
                     /rpt_family="L1PA13_5end"
     exon            complement(12550. .12611)
                     /note="MZEF"
                     /evidence=not_experimental
     exon            complement(12810. .12873)
                     /note="MZEF"
                     /evidence=not_experimental
     exon            13577. .13759
                     /note="MZEF"
                     /evidence=not_experimental
     repeat_region   14607. .14633
                     /note="AT_rich"
                     /rpt_family="Low_complexity"
     repeat_region   complement(15088. .15448)
                     /rpt_family="THE1B"
     exon            15178. .15330
                     /note="GRAIL"
                     /evidence=not_experimental
     repeat_region   complement(15947. .15997)
                     /rpt_family="L1MA9"
     repeat_region   16433. .16457
                     /note="AT_rich"
                     /rpt_family="Low_complexity"
     repeat_region   complement(17199. .17507)
                     /rpt_family="L2a"
     exon            17452. .17562
                     /note="GENSCAN"
     repeat_region   complement(17506. .18111)
                     /rpt_family="L2"
     exon            17795. .18032
                     /note="GRAIL"
                     /evidence=not_experimental
     repeat_region   18238. .18500
                     /rpt_family="MLT1H"
     repeat_region   complement(18825. .18851)
                     /note="AT_rich"
                     /rpt_family="Low_complexity"
     repeat_region   complement(18878. .18898)
                     /note="AT_rich"
                     /rpt_family="Low_complexity"
     repeat_region   complement(18938. .18977)
                     /note="AT_rich"
                     /rpt_family="Low_complexity"
     exon            complement(19054. .19199)
                     /note="GRAIL"
```

```
                /evidence=not_experimental
exon            19154. .19261
                /note="GENSCAN"
exon            complement(20188. .20326)
                /note="GRAIL"
                /evidence=not_experimental
repeat_region   complement(20689. .20863)
                /rpt_family="MER5A"
repeat_region   complement(21353. .21379)
                /note="AT_rich"
                /rpt_family="Low_complexity"
exon            21378. .21447
                /note="GRAIL"
                /evidence=not_experimental
repeat_region   complement(21509. .21667)
                /rpt_family="AluSg1"
exon            complement(21638. .21785)
                /note="MZEF"
                /evidence=not_experimental
repeat_region   22368. .22902
                /rpt_family="MLT1E"
repeat_region   24714. .25090
                /rpt_family="MER57B"
repeat_region   complement(25217. .25525)
                /rpt_family="AluSc"
exon            25801. .25905
                /note="MZEF"
                /evidence=not_experimental
repeat_region   25807. .26115
                /rpt_family="AluY"
exon            complement(26760. .26806)
                /note="MZEF"
                /evidence=not_experimental
repeat_region   complement(27112. .27454)
                /note="(GA)n"
                /rpt_family="Simple_repeat"
repeat_region   complement(27132. .27443)
                /rpt_family="AluSp"
exon            28969. .29151
                /note="GRAIL"
                /evidence=not_experimental
repeat_region   complement(29179. .29264)
                /rpt_family="L1PA8"
repeat_region   29727. .30189
                /rpt_family="MER31A"
repeat_region   30395. .30522
                /rpt_family="MIR"
exon            30759. .30834
                /note="GRAIL"
                /evidence=not_experimental
repeat_region   31435. .31607
                /note="AT_rich"
                /rpt_family="Low_complexity"
repeat_region   complement(31610. .31678)
                /note="(TA)n"
                /rpt_family="Simple_repeat"
exon            complement(31923. .31969)
                /note="XPOUND"
                /evidence=not_experimental
exon            complement(33187. .33252)
                /note="GRAIL"
                /evidence=not_experimental
repeat_region   complement(33838. .34147)
                /rpt_family="AluSc"
repeat_region   complement(34805. .34853)
                /rpt_family="AluSg"
repeat_region   35084. .35200
                /rpt_family="L1MC5"
repeat_region   complement(35212. .35730)
                /rpt_family="L1"
exon            complement(35238. .35438)
                /note="GRAIL"
                /evidence=not_experimental
exon            complement(35238. .35315)
                /note="MZEF"
                /evidence=not_experimental
exon            complement(35326. .35375)
                /note="XPOUND"
                /evidence=not_experimental
exon            complement(35438. .35682)
                /note="GENSCAN"
exon            complement(35438. .35682)
                /note="GRAIL"
                /evidence=not_experimental
repeat_region   35725. .36564
                /rpt_family="L1"

  Query Match              2.4%;  Score 21;  DB 9;  Length 157270;
  Best Local Similarity   100.0%;  Pred. No. 8.6;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        346 TTAACTAAAGCATTTGAACTT 366
              |||||||||||||||||||||
Db      81200 TTAACTAAAGCATTTGAACTT 81220


Search completed: May 30, 2003, 19:07:51
Job time : 2487 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 30, 2003, 17:23:55 ; Search time 280 Seconds
(without alignments)
6924.895 Million cell updates/sec

Title: US-09-925-637-63
Perfect score: 861
Sequence: 1 atgacgaatctaccgatgaa.....tattagaaatcgttgattta 861

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

rd size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	861	100.0	861	22	AAS00821	S. aureus HGS072 e
c	2	807	93.7	1893	18	AAV74466	Staphylococcus aur
	3	606	70.4	882	21	AAA92031	Staphylococcus aur
	4	489	56.8	882	23	AAS54848	Staphylococcus aur
	5	474	55.1	864	23	AAS51602	Staphylococcus aur
c	6	413	48.0	413	23	AAS50245	Staphylococcus aur
c	7	337	39.1	337	23	AAS50698	Staphylococcus aur
	8	25	2.9	783	22	AAH52415	S. epidermidis ope
	9	25	2.9	909	24	ABN90871	Staphylococcus epi
	10	25	2.9	3444	22	AAH54780	S. epidermidis gen
c	11	25	2.9	3454	22	AAH54634	S. epidermidis gen
c	12	25	2.9	4045	22	AAH54565	S. epidermidis gen
	13	20	2.3	2365589	24	ABA90521	Genomic sequence o
c	14	19	2.2	151	22	ABA74941	Human foetal liver
c	15	19	2.2	151	22	ABA39630	Probe #18096 for g
c	16	19	2.2	151	22	AAK23438	Human brain expres
c	17	19	2.2	151	22	AAK49586	Human bone marrow
c	18	19	2.2	151	22	AAI26694	Probe #16627 for g
c	19	19	2.2	151	22	AAI55466	Probe #24152 used
c	20	19	2.2	151	24	ABS23069	Human genome-deriv
c	21	19	2.2	600	22	ABA62413	Human foetal liver
c	22	19	2.2	600	22	ABA29746	Probe #8212 for ge
c	23	19	2.2	600	22	AAK10751	Human brain expres
c	24	19	2.2	600	22	AAK36630	Human bone marrow
c	25	19	2.2	600	22	AAI17482	Probe #7415 for ge
c	26	19	2.2	600	22	AAI42391	Probe #11077 used
c	27	19	2.2	600	24	ABS10626	Human genome-deriv
	28	19	2.2	809	24	ABQ33820	Oligonucleotide fo
c	29	19	2.2	809	24	ABQ33821	Oligonucleotide fo
	30	19	2.2	873	24	ABN66184	Streptococcus poly
c	31	19	2.2	6072	23	ABL07772	Drosophila melanog
c	32	19	2.2	640681	24	ABA92787	Buchnera sp. genom
	33	19	2.2	2155561	24	ABN71527	Streptococcus poly
c	34	18	2.1	60	19	AAV45107	SDF-1-alpha primer
	35	18	2.1	118	22	AAS30093	Human lung antigen
	36	18	2.1	164	21	AAC07133	Human secreted pro
c	37	18	2.1	322	20	AAV87896	EST clone FG99. H
	38	18	2.1	350	23	ABV50402	Human prostate exp
	39	18	2.1	491	24	ABN61915	Human cancer relat
	40	18	2.1	588	22	AAH12725	Human cDNA clone (
c	41	18	2.1	721	22	AAS27074	cDNA encoding nove
c	42	18	2.1	721	23	ABK43767	DNA encoding novel
	43	18	2.1	919	21	AAC56354	Pinus radiata tran
	44	18	2.1	1301	12	AAQ14532	T cell protein 19.
	45	18	2.1	1301	15	AAQ63615	T-cell transmembra

ALIGNMENTS

RESULT 1
AAS00821
ID AAS00821 standard; DNA; 861 BP.
XX
AC AAS00821;
XX
DT 04-JUL-2001 (first entry)
XX
DE S. aureus HGS072 encoding Farnesyl diphosphatesynthase, IspA.
XX
KW Farnesyl diphosphatesynthase; IspA; immunogen; vaccine; antibody;
KW wound infection; cellulitis; burn infection; eyelid infection;
KW food poisoning; joint infection; neonatal conjunctivitis; osteomyelitis;
KW skin infection; scalded skin syndrome; toxic epidermal necrosis;
KW Ritter's disease; Lyell's disease; toxic shock syndrome; endocarditis;
KW HGS072; ds.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT CDS 1..861
FT /*tag= a
FT /product= "IspA"
FT /partial
FT /note= "No stop codon"
XX
PN WO200116292-A2.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-US23773.
XX

```
PR   01-SEP-1999;   99US-0151933.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Choi GH;
XX
DR   WPI; 2001-183259/18.
DR   P-PSDB; AAU00859.
XX
PT   New isolated nucleic acid for use in diagnosing Staphylococcus
PT   infections and in vaccines for eliciting immune responses to the
PT   infections -
XX
PS   Claim 1; Page 23; 225pp; English.
XX
CC   The sequence encodes S. aureus IspA (Farnesyl diphosphatesynthase).
CC   The polynucleotides of the invention are used to detect Staphylococcus
CC   nucleic acids in a biological sample from an animal for diagnosing
CC   Staphylococcus infections. The polypeptides of the invention are used to
     detect anti-Staphylococcus antibodies in a biological sample from an
     animal to diagnose Staphylococcus infections. The polypeptides are also
     used in vaccines to elicit protective antibodies in an animal to a member
CC   of the Staphylococcus genus and for preventing or attenuating an
CC   infection caused by a member of the Staphylococcus genus e.g wound
CC   infection, cellulitis, burn infection, eyelid infection, food poisoning,
CC   joint infection, neonatal conjunctivitis, osteomyelitis, skin infection,
CC   scalded skin syndrome (also known as toxic epidermal necrosis, Ritter's
CC   disease and Lyell's disease), toxic shock syndrome and endocarditis. The
CC   polynucleotides may also be used in vaccines and for preventing or
CC   attenuating a Staphylococcus infection. Antibodies to the polypeptides
CC   may be used to purify, detect and target the polypeptides in in vitro and
CC   in vivo diagnostic and therapeutic methods.
XX
SQ   Sequence 861 BP; 323 A; 114 C; 180 G; 244 T; 0 other;

  Query Match            100.0%;  Score 861;  DB 22;  Length 861;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  861;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||||||||||||||||||||
Db        841 TTATTAGAAATCGTTGATTTA 861


RESULT 2
AAV74466/c
ID   AAV74466 standard; DNA; 1893 BP.
XX
AC   AAV74466;
XX
DT   16-MAR-1999  (first entry)
XX
DE   Staphylococcus aureus contig SEQ ID #155.
XX
KW   Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW   cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW   skin infection; surgical wound infection; scalded skin syndrome;
KW   toxic shock syndrome; ds.
XX
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1261..1320
FT                   /*tag= a
FT                   /note= "these bases represent a line of missing text in
FT                         the sequence listing in the specification. They
FT                         are included to maintain the nucleotide numbering
FT                         given in the specification for this DNA sequence"
XX
PN   EP786519-A2.
XX
PD   30-JUL-1997.
XX
PF   07-JAN-1997;   97EP-0100117.
XX
PR   05-JAN-1996;   96US-0009861.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Barash SC,  Choi GH,  Dillon PJ,  Fannon MR,  Kunsch CA;
PI   Rosen CA;
XX
DR   WPI; 1997-374922/35.
XX
PT   Polynucleotide(s) and proteins derived from Staphylococcus aureus -
PT   stored on computer readable medium and used in the production of
PT   anti-S.aureus vaccines
XX
PS   Claim 1; Page 784-785; 3271pp; English.
XX
```

CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the S.aureus DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against S.aureus infection. The
CC polypeptides can also be used in a kit for the immunodetection of
CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC skin and surgical wound infections, scalded skin syndrome, toxic shock
CC syndrome, etc. Organisms transformed with the DNA sequences can be used
CC for recombinant production of the polypeptides. The new DNA sequences
CC (and their fragments) are useful as primers or probes for isolating
CC homologues of any of the S.aureus DNA sequences contained on the
CC computer readable medium.
XX
SQ Sequence 1893 BP; 552 A; 336 C; 233 G; 708 T; 64 other;

Query Match 93.7%; Score 807; DB 18; Length 1893;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 807; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        55 GTTGCGATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTA 114
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1260 GTTGCGATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTA 1201

Qy       115 AATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAAT 174
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1200 AATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAAT 1141

Qy       175 ACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCA 234
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1140 ACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCA 1081

Qy       235 CTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACA 294
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1080 CTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACA 1021

Qy       295 AATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAA 354
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1020 AATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAA 961

Qy       355 GCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTA 414
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       960 GCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTA 901

Qy       415 CAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATG 474
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       900 CAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATG 841

Qy       475 CAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACA 534
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       840 CAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACA 781

Qy       535 GGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACA 594
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       780 GGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACA 721

Qy       595 ACTAAAGAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGAT 654
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       720 ACTAAAGAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGAT 661

Qy       655 GATTTATTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTT 714
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       660 GATTTATTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTT 601

Qy       715 GAAAATAATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTG 774
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       600 GAAAATAATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTG 541
```

```
Qy       775 ACTTATCATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACA 834
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       540 ACTTATCATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACA 481

Qy       835 AAACACTTATTAGAAATCGTTGATTTA 861
             |||||||||||||||||||||||||||
Db       480 AAACACTTATTAGAAATCGTTGATTTA 454
```

RESULT 3
AAA92031
ID AAA92031 standard; DNA; 882 BP.
XX
AC AAA92031;
XX
DT 12-JAN-2001 (first entry)
XX
DE Staphylococcus aureus ispA coding sequence.
XX
KW IspA; bacterial disease; respiratory tract infection;
KW gastrointestinal infection; cardiac infection; Helicobacter pylori;
KW stomach cancer; stomach ulcer; gastritis; ds.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT CDS 1..882
FT /*tag= a
FT /product= "IspA"
XX
PN US6107058-A.
XX
PD 22-AUG-2000.
XX
PF 26-MAR-1999; 99US-0276873.
XX
PR 26-MAR-1999; 99US-0276873.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Gwynn M, Wilding EI;
XX
DR WPI; 2000-578535/54.
DR P-PSDB; AAB23333.
XX
PT Novel farnesyl diphosphate synthase polynucleotide from staphylococcus
PT aureus useful for diagnosis and treatment of bacterial infections and
PT as hybridization probe for isolating genomic clones -
XX
PS Claim 1; column 1-4; 15pp; English.
XX
CC The present sequence is the coding sequence for the Staphylococcus aureus
CC IspA protein. This gene and the protein it encodes can be used in many
CC research assays, as well as treatments for bacterial diseases such as
CC infections of the respiratory tract (including otitis media, bacterial
CC tracheitis, acute epiglottitis, thyroiditis, empyema and lung abscesses),
CC cardiac infections such as infective endocarditis, gastrointestinal
CC infections including secretory diarrhoea, splenic abscesses and
CC retroperitoneal abscesses, CNS infections such as cerebral abscesses, eye
CC infections (including blepharitis, conjunctivitis, keratitis,
CC endophthalmitis, preseptal and orbital cellulitis and darcryocystitis),
CC kidney and urinary tract infections such as epididymitis, intrarenal and
CC perinephric abscesses and toxic shock syndrome, skin diseases (including
CC impetigo, folliculitis, cutaneous abscesses, wound infection and
CC bacterial myositis), bone and joint infections such as septic
CC arthritis and osteomyelitis, septic thrombophlebitis, food poisoning and
CC scalded skin syndrome. In addition, they can be used to treat
CC diseases caused by Helicobacter pylori, including stomach cancer, stomach
CC ulcers and gastritis.
XX
SQ Sequence 882 BP; 328 A; 118 C; 186 G; 250 T; 0 other;

Query Match 70.4%; Score 606; DB 21; Length 882;

```
  Best Local Similarity   99.4%;  Pred. No. 4.6e-291;
  Matches  856;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGCATGTTGTATTCATTAAATGCT 120

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTGTGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

          241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              |||||||||||||||||||||||||||||||||||| | ||||||||||||||||||||
          241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAAGGAAATTAACAAATCAT 300

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAA 600

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||||||||||||||||||||
Db        841 TTATTAGAAATCGTTGATTTA 861


RESULT 4
AAS54848
ID   AAS54848 standard; DNA; 882 BP.
XX
AC   AAS54848;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus DNA for cellular proliferation protein #1160.
XX
KW   Antisense; ds; prokaryotic cellular proliferation gene;
KW   antibiotic; antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   P-PSDB; AAU36989.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Claim 27; Seq ID No 8485; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence encodes an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 882 BP; 330 A; 116 C; 185 G; 251 T; 0 other;

  Query Match             56.8%;  Score 489;  DB 23;  Length 882;
  Best Local Similarity   99.8%;  Pred. No. 6.5e-233;
  Matches  539;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCAT 300

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
```

```
RESULT 5
AAS51602
ID   AAS51602 standard; DNA; 864 BP.
XX
AC   AAS51602;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus DNA for cellular proliferation protein #19.
XX
KW   Antisense; ds; prokaryotic cellular proliferation gene;
KW   antibiotic; antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   P-PSDB; AAU33743.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Claim 27; Seq ID No 4184; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence encodes an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 864 BP; 323 A; 112 C; 182 G; 247 T; 0 other;

  Query Match            55.1%;  Score 474;  DB 23;  Length 864;
  Best Local Similarity  99.8%;  Pred. No. 1.9e-225;
  Matches  524;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         16 ATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTA 60

Qy         76 ATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATC 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATC 120

Qy        136 CGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATG 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATG 180

Qy        196 AAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCA 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCA 240

Qy        256 GCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAG 315
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        241 GCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCATAAAGTATATGGTGAG 300

Qy        316 TGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGT 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGT 360

Qy        376 GATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGT 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGT 420

Qy        436 GGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATT 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATT 480

Qy        496 GATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              |||||||||||||||||||||||||||||||||||||||||||||
Db        481 GATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 525
```

```
RESULT 6
AAS50245/c
ID   AAS50245 standard; DNA; 413 BP.
XX
AC   AAS50245;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus cellular proliferation inhibitory sequence #1469.
XX
KW   Antisense; ss; prokaryotic cellular proliferation;
KW   antibiotic; antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
```

```
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
XX
PT   New polynucleotides for the identification and development of
     antibiotics, comprise sequences of antisense nucleic acids -

     Claim 1; Seq ID No 2822; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence is an antisense
CC   oligonucleotide of the invention.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 413 BP; 115 A; 89 C; 58 G; 151 T; 0 other;

  Query Match            48.0%;  Score 413;  DB 23;  Length 413;
  Best Local Similarity  100.0%;  Pred. No. 3.9e-195;
  Matches  413;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        160 TTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATG 219
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 TTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATG 354

Qy        220 ATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGA 279
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 ATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGA 294

Qy        280 CGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGAT 339
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 CGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGAT 234

Qy        340 GCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTA 399
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 GCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTA 174

Qy        400 AAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGT 459
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 AAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGT 114

Qy        460 CAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATA 519
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 CAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATA 54

Qy        520 CACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCAGA 572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         53 CACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCAGA 1


RESULT 7
AAS50698/c
ID   AAS50698 standard; DNA; 337 BP.
XX
AC   AAS50698;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus cellular proliferation inhibitory sequence #1922.
XX
KW   Antisense; ss; prokaryotic cellular proliferation;
KW   antibiotic; antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Claim 1; Seq ID No 3275; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence is an antisense
CC   oligonucleotide of the invention.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 337 BP; 96 A; 71 C; 47 G; 123 T; 0 other;

  Query Match            39.1%;  Score 337;  DB 23;  Length 337;
  Best Local Similarity  100.0%;  Pred. No. 2.4e-157;
  Matches  337;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        225 TACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGG 284
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 TACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGG 278

Qy        285 AAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTT 344
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 AAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTT 218

Qy        345 ATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAAT 404
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 ATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAAT 158

Qy        405 AAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAAT 464
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 AAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAAT 98

Qy        465 GTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAA 524
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 GTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAA 38

Qy        525 AACAAAAACAGGAGCATTATTAACTTTTGCGGTTATG 561
                |||||||||||||||||||||||||||||||||||
Db         37 AACAAAAACAGGAGCATTATTAACTTTTGCGGTTATG 1
```

RESULT 8
AAH52415
ID AAH52415 standard; DNA; 783 BP.
XX
AC AAH52415;
XX
DT 03-SEP-2001 (first entry)
XX
DE S. epidermidis open reading frame nucleotide sequence SEQ ID NO:223.
XX
KW Staphylococcus epidermidis SR1 strain; infection; diagnosis;
KW vaccination; endocarditis; ds.
XX
OS Staphylococcus epidermidis.
XX
PN WO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US30782.
XX
PR 09-NOV-1999; 99US-0164258.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Kimmerly WJ;
XX
DR WPI; 2001-316495/33.
DR P-PSDB; AAG81565.
XX
PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis -
XX
PS Claim 8; Page 103; 2188pp; English.
XX
CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ Sequence 783 BP; 297 A; 93 C; 147 G; 246 T; 0 other;

```
  Query Match             2.9%;  Score 25;  DB 22;  Length 783;
  Best Local Similarity   100.0%;  Pred. No. 0.028;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        281 GAGGAAAATTAACAAATCATAAAGT 305
                |||||||||||||||||||||||||
Db        182 GAGGAAAATTAACAAATCATAAAGT 206
```

RESULT 9
ABN90871
ID ABN90871 standard; DNA; 909 BP.
XX
AC ABN90871;
XX
DT 24-JUL-2002 (first entry)
XX
DE Staphylococcus epidermidis ORF nucleic acid sequence SEQ ID NO:334.
XX
KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW antibacterial; gene therapy; gene; ds.
XX
OS Staphylococcus epidermidis.
XX
PN US6380370-B1.
XX
PD 30-APR-2002.
XX
PF 13-AUG-1998; 98US-0134001.
XX
PR 14-AUG-1997; 97US-055779P.
PR 08-NOV-1997; 97US-064964P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Doucette-Stamm LA, Bush D;
XX
DR WPI; 2002-381255/41.
DR P-PSDB; ABP38326.
XX
PT Novel isolated nucleic acid encoding a Staphylococcus epidermidis
PT polypeptide, useful for diagnosing and treating bacterial infections -
XX
PS Disclosure; SEQ ID 334; 267pp; English.
XX
CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC antibacterial activity and can be used in gene therapy. The sequences
CC can also be used in the diagnosis and treatment of bacterial infections,
CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life
CC cycle or inhibit S. epidermidis infection.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site.
XX
SQ Sequence 909 BP; 361 A; 106 C; 165 G; 277 T; 0 other;

```
  Query Match             2.9%;  Score 25;  DB 24;  Length 909;
  Best Local Similarity   100.0%;  Pred. No. 0.028;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        281 GAGGAAAATTAACAAATCATAAAGT 305
```

```
                    ||||||||||||||||||||||||||
Db        308 GAGGAAAATTAACAAATCATAAAGT 332


RESULT 10
AAH54780
ID   AAH54780 standard; DNA; 3444 BP.
XX
AC   AAH54780;
XX
DT   03-SEP-2001  (first entry)
XX
DE   S. epidermidis genomic polynucleotide sequence SEQ ID NO:4144.
XX
KW   Staphylococcus epidermidis SR1 strain; infection; diagnosis;
KW   vaccination; endocarditis; ds.
XX
OS   Staphylococcus epidermidis.
XX
     WO200134809-A2.

     17-MAY-2001.
XX
PF   09-NOV-2000; 2000WO-US30782.
XX
PR   09-NOV-1999;   99US-0164258.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Kimmerly WJ;
XX
DR   WPI; 2001-316495/33.
XX
PT   Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT   useful for vaccinating against infections, e.g. endocarditis -
XX
PS   Claim 8; Page 1844-1845; 2188pp; English.
XX
CC   AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC   (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC   (I) and (II) can have antibacterial activity and therefore can be used
CC   in vaccination. The nucleic acids (I) may be used to produce the
CC   S. epidermidis polypeptides (II) via the production of vectors
CC   containing them which are used to produce hosts cells which express the
CC   polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC   used to vaccinate subjects and to raise antibodies against the bacteria.
CC   The polypeptides may also be used to assay for other inhibitors of their
CC   activity and therefore identify compounds that may be used for the
CC   treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC   AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC   polynucleotide sequences from the present invention. AAH55091 to
CC   AAH55098 represent oligonucleotide sequences and primers which are used
CC   in the exemplification of the present invention.
CC   N.B. The present invention specifically claims all the polynucleotide
CC   sequences given in the sequence listing of the present specification,
CC   however the sequence listing only goes up to SEQ ID NO:4454 so even
CC   though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC   no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ   Sequence 3444 BP; 1365 A; 451 C; 571 G; 1057 T; 0 other;

  Query Match             2.9%;  Score 25;  DB 22;  Length 3444;
  Best Local Similarity   100.0%;  Pred. No. 0.028;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        281 GAGGAAAATTAACAAATCATAAAGT 305
              |||||||||||||||||||||||||
Db       1046 GAGGAAAATTAACAAATCATAAAGT 1070


RESULT 11
AAH54634/c
ID   AAH54634 standard; DNA; 3454 BP.
XX
AC   AAH54634;
XX
DT   03-SEP-2001  (first entry)
XX
DE   S. epidermidis genomic polynucleotide sequence SEQ ID NO:3998.
XX
KW   Staphylococcus epidermidis SR1 strain; infection; diagnosis;
KW   vaccination; endocarditis; ds.
XX
OS   Staphylococcus epidermidis.
XX
PN   WO200134809-A2.
XX
PD   17-MAY-2001.
XX
PF   09-NOV-2000; 2000WO-US30782.
XX
PR   09-NOV-1999;   99US-0164258.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Kimmerly WJ;
XX
DR   WPI; 2001-316495/33.
XX
PT   Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT   useful for vaccinating against infections, e.g. endocarditis -
XX
PS   Claim 8; Page 1671-1672; 2188pp; English.
XX
CC   AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC   (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC   (I) and (II) can have antibacterial activity and therefore can be used
CC   in vaccination. The nucleic acids (I) may be used to produce the
CC   S. epidermidis polypeptides (II) via the production of vectors
CC   containing them which are used to produce hosts cells which express the
CC   polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC   used to vaccinate subjects and to raise antibodies against the bacteria.
CC   The polypeptides may also be used to assay for other inhibitors of their
CC   activity and therefore identify compounds that may be used for the
CC   treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC   AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC   polynucleotide sequences from the present invention. AAH55091 to
CC   AAH55098 represent oligonucleotide sequences and primers which are used
CC   in the exemplification of the present invention.
CC   N.B. The present invention specifically claims all the polynucleotide
CC   sequences given in the sequence listing of the present specification,
CC   however the sequence listing only goes up to SEQ ID NO:4454 so even
CC   though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC   no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ   Sequence 3454 BP; 1005 A; 574 C; 493 G; 1382 T; 0 other;

  Query Match             2.9%;  Score 25;  DB 22;  Length 3454;
  Best Local Similarity   100.0%;  Pred. No. 0.028;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        281 GAGGAAAATTAACAAATCATAAAGT 305
              |||||||||||||||||||||||||
Db         77 GAGGAAAATTAACAAATCATAAAGT 53


RESULT 12
AAH54565/c
ID   AAH54565 standard; DNA; 4045 BP.
XX
AC   AAH54565;
XX
DT   03-SEP-2001  (first entry)
XX
DE   S. epidermidis genomic polynucleotide sequence SEQ ID NO:3929.
XX
```

KW Staphylococcus epidermidis SR1 strain; infection; diagnosis;
KW vaccination; endocarditis; ds.
XX
OS Staphylococcus epidermidis.
XX
PN WO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US30782.
XX
PR 09-NOV-1999; 99US-0164258.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Kimmerly WJ;
XX
DR WPI; 2001-316495/33.
XX
PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis, useful for vaccinating against infections, e.g. endocarditis -
XX
PS Claim 8; Page 1586-1588; 2188pp; English.
XX
CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ Sequence 4045 BP; 1249 A; 691 C; 536 G; 1569 T; 0 other;

Query Match 2.9%; Score 25; DB 22; Length 4045;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 281 GAGGAAAATTAACAAATCATAAAGT 305
||||||||||||||||||||||||||
Db 2033 GAGGAAAATTAACAAATCATAAAGT 2009

RESULT 13
ABA90521
ID ABA90521 standard; DNA; 2365589 BP.
XX
AC ABA90521;
XX
DT 16-MAY-2002 (first entry)
XX
DE Genomic sequence of Lactococcus lactis IL1403.
XX
KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese; ds.
XX
OS Lactococcus lactis IL1403.
XX
PN FR2807446-A1.
XX
PD 12-OCT-2001.
XX
PF 11-APR-2000; 2000FR-0004630.
XX
PR 11-APR-2000; 2000FR-0004630.
XX
PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
XX
PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;
XX
DR WPI; 2002-043418/06.
XX
PT New nucleotide sequence useful in the identification or Lactococcus
PT lactis and related species -
XX
PS Claim 1; SEQ ID 1; 2504pp; French.
XX
CC The present invention is related to a Lactococcus lactis nucleotide
CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The
CC nucleic acid sequence is useful in the detection and/or amplification of
CC nucleic acid sequence, particularly to identify Lactococcus lactis or
CC related species. The proteins of the invention are useful for the
CC biosynthesis or biodegradation of a composition of interest. The
CC invention helps research in lactic bacteria, particularly useful in the
CC production of yogurt and cheese.
CC Note: The sequence data for this patent is based on equivalent patent
CC WO200177334 (published 18-OCT-2001) which is available in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2365589 BP; 765914 A; 415261 C; 420487 G; 763927 T; 0 other;

Query Match 2.3%; Score 20; DB 24; Length 2365589;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 259 ATGGATAATGATGATTATCG 278
||||||||||||||||||||
Db 881007 ATGGATAATGATGATTATCG 881026

RESULT 14
ABA74941/c
ID ABA74941 standard; DNA; 151 BP.
XX
AC ABA74941;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #23246.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00669.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX

PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
PS Claim 4; SEQ ID NO 23246; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 151 BP; 31 A; 22 C; 24 G; 74 T; 0 other;

Query Match 2.2%; Score 19; DB 22; Length 151;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 503 AAACTTTGGAAATGATACA 521
||||||||||||||||||||
Db 139 AAACTTTGGAAATGATACA 121

RESULT 15
ABA39630/c
ID ABA39630 standard; DNA; 151 BP.
XX
AC ABA39630;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #18096 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
XX
PS Claim 4; SEQ ID No 18096; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 151 BP; 31 A; 22 C; 24 G; 74 T; 0 other;

Query Match 2.2%; Score 19; DB 22; Length 151;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 503 AAACTTTGGAAATGATACA 521
||||||||||||||||||||
Db 139 AAACTTTGGAAATGATACA 121

Search completed: May 30, 2003, 18:26:12
Job time : 283 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 30, 2003, 17:36:25 ; Search time 61 Seconds
(without alignments)
4328.667 Million cell updates/sec

Title: US-09-925-637-63
Perfect score: 861
Sequence: 1 atgacgaatctaccgatgaa.....tattagaaatcgttgattta 861

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	606	70.4	882	3	US-09-276-873-1	Sequence 1, Appli
	2	25	2.9	909	4	US-09-134-001C-334	Sequence 334, App
	3	18	2.1	763	4	US-09-484-970B-110	Sequence 110, App
	4	18	2.1	1301	1	US-07-686-322A-1	Sequence 1, Appli
	5	18	2.1	1301	1	US-08-002-999-1	Sequence 1, Appli
	6	18	2.1	5163	3	US-08-700-651-1	Sequence 1, Appli
	7	18	2.1	5163	3	US-08-928-361B-4	Sequence 4, Appli
	8	18	2.1	5318	3	US-08-700-651-2	Sequence 2, Appli
	9	18	2.1	5318	3	US-08-928-361B-3	Sequence 3, Appli
	10	17	2.0	604	4	US-09-385-982-404	Sequence 404, App
	11	17	2.0	616	4	US-09-385-982-173	Sequence 173, App
	12	17	2.0	660	2	US-08-902-516-1	Sequence 1, Appli
	13	17	2.0	1164	4	US-09-134-001C-2199	Sequence 2199, Ap
	14	17	2.0	1511	2	US-08-809-763-3	Sequence 3, Appli
	15	17	2.0	1511	3	US-08-956-253-3	Sequence 3, Appli
	16	17	2.0	2335	2	US-08-300-584-3	Sequence 3, Appli
	17	17	2.0	2335	4	US-08-476-123-3	Sequence 3, Appli
	18	17	2.0	4656	4	US-09-150-460B-4	Sequence 4, Appli
	19	17	2.0	6614	4	US-09-150-460B-3	Sequence 3, Appli
	20	17	2.0	7358	4	US-09-058-483-8	Sequence 8, Appli
c	21	17	2.0	4403765	4	US-09-103-840A-2	Sequence 2, Appli
c	22	17	2.0	4411529	4	US-09-103-840A-1	Sequence 1, Appli
	23	16	1.9	90	4	US-09-507-819-7	Sequence 7, Appli
c	24	16	1.9	384	4	US-09-438-906-25	Sequence 25, Appl
c	25	16	1.9	507	4	US-09-134-001C-2678	Sequence 2678, Ap
	26	16	1.9	621	4	US-09-134-001C-1593	Sequence 1593, Ap
	27	16	1.9	690	4	US-09-134-001C-2192	Sequence 2192, Ap
	28	16	1.9	732	4	US-09-134-001C-1402	Sequence 1402, Ap
	29	16	1.9	861	3	US-09-187-050-4	Sequence 4, Appli
	30	16	1.9	863	1	US-08-592-126-71	Sequence 71, Appl
	31	16	1.9	885	3	US-09-187-050-11	Sequence 11, Appl
	32	16	1.9	909	2	US-08-286-819A-22	Sequence 22, Appl
	33	16	1.9	909	3	US-08-980-357-22	Sequence 22, Appl
	34	16	1.9	980	4	US-09-171-209-8	Sequence 8, Appli
	35	16	1.9	1179	3	US-09-187-050-13	Sequence 13, Appl
	36	16	1.9	1179	3	US-09-187-050-15	Sequence 15, Appl
	37	16	1.9	1179	3	US-09-187-050-17	Sequence 17, Appl
	38	16	1.9	1179	3	US-09-187-050-19	Sequence 19, Appl
	39	16	1.9	1179	3	US-09-187-050-21	Sequence 21, Appl
	40	16	1.9	1179	3	US-09-187-050-23	Sequence 23, Appl
	41	16	1.9	1179	3	US-09-187-050-25	Sequence 25, Appl
	42	16	1.9	1378	1	US-08-592-126-72	Sequence 72, Appl
c	43	16	1.9	1508	3	US-08-714-918-31	Sequence 31, Appl
c	44	16	1.9	1508	4	US-09-265-315-31	Sequence 31, Appl
c	45	16	1.9	1508	4	US-09-265-315-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1
US-09-276-873-1
; Sequence 1, Application US/09276873
; Patent No. 6107058
; GENERAL INFORMATION:
; APPLICANT: Wilding, Edwina Imogen
; APPLICANT: Gwynn, Michael
; TITLE OF INVENTION: IspA
; FILE REFERENCE: GM10208
; CURRENT APPLICATION NUMBER: US/09/276,873
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 882
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-276-873-1

Query Match 70.4%; Score 606; DB 3; Length 882;
Best Local Similarity 99.4%; Pred. No. 2.7e-293;
Matches 856; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy       1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy      61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db      61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGCATGTTGTATTCATTAAATGCT 120

Qy     121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy     181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
           |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TATGAGTTAGGTGTGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy     241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
           |||||||||||||||||||||||||||||||||||| | ||||||||||||||||||||
Db     241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAAGGAAATTAACAAATCAT 300

Qy     301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy     361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
```

```
Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAA 600

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||||||||||||||||||||
Db        841 TTATTAGAAATCGTTGATTTA 861
```

RESULT 2
US-09-134-001C-334
; Sequence 334, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 334
; LENGTH: 909
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-334

Query Match 2.9%; Score 25; DB 4; Length 909;
Best Local Similarity 100.0%; Pred. No. 0.0036;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        281 GAGGAAAATTAACAAATCATAAAGT 305
              |||||||||||||||||||||||||
Db        308 GAGGAAAATTAACAAATCATAAAGT 332
```

RESULT 3
US-09-484-970B-110
; Sequence 110, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmuth, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 110
; LENGTH: 763
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6426186 009051.5CB1
US-09-484-970B-110

Query Match 2.1%; Score 18; DB 4; Length 763;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         15 GATGAATAAATTAATAGA 32
              ||||||||||||||||||
Db        212 GATGAATAAATTAATAGA 229
```

RESULT 4
US-07-686-322A-1
; Sequence 1, Application US/07686322A
; Patent No. 5312733
; GENERAL INFORMATION:
; APPLICANT: MacLeod Dr., Carol L.
; TITLE OF INVENTION: No. 5312733el T-cell Lymphoma cDNA Clones
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patent Department, Fulbright & Jaworski
; STREET: 1301 McKinney, Suit 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/686,322A
; FILING DATE: 19910411
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/509684
; FILING DATE: 13-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene A.
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: D-5232-CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 651-3634
; TELEFAX: (713) 651-5246
; TELEX: Western Union 762829
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1301 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Mouse
; STRAIN: AKR1 Jackson
; INDIVIDUAL ISOLATE: SL12 cell line
; TISSUE TYPE: Lymphoma
; CELL TYPE: T-cell
; CELL LINE: SL12.3 and SL12.4

```
;    IMMEDIATE SOURCE:
;      CLONE:  19.1
US-07-686-322A-1

  Query Match             2.1%;  Score 18;  DB 1;  Length 1301;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        828 CAATACAAAACACTTATT 845
              ||||||||||||||||||
Db          3 CAATACAAAACACTTATT 20


RESULT 5
US-08-002-999-1
; Sequence 1, Application US/08002999
; Patent No. 5440017
;  GENERAL INFORMATION:
;    APPLICANT:  MacLeod Dr., Carol L.
     TITLE OF INVENTION:  No. 5440017el T-cell Lymphoma cDNA Clones
     NUMBER OF SEQUENCES:  6
     CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patent Department, Fulbright & Jaworski
;      STREET:  1301 McKinney, Suite 5100
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77010-3095
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/002,999
;      FILING DATE:  19930111
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/686,322
;      FILING DATE:  11-APR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Adler, Benjamin A.
;      REGISTRATION NUMBER:  35,423
;      REFERENCE/DOCKET NUMBER:  D-5232-DIV
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (713) 651-5587
;      TELEFAX:  (713) 651-5246
;      TELEX:  Western Union 762829
   INFORMATION FOR SEQ ID NO:  1:
     SEQUENCE CHARACTERISTICS:
;      LENGTH:  1301 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Mouse
;      STRAIN:  AKR1 Jackson
;      INDIVIDUAL ISOLATE:  SL12 cell line
;      TISSUE TYPE:  Lymphoma
;      CELL TYPE:  T-cell
;      CELL LINE:  SL12.3 and SL12.4
;    IMMEDIATE SOURCE:
;      CLONE:  19.1
US-08-002-999-1

  Query Match             2.1%;  Score 18;  DB 1;  Length 1301;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        828 CAATACAAAACACTTATT 845
              ||||||||||||||||||
Db          3 CAATACAAAACACTTATT 20


RESULT 6
US-08-700-651-1
; Sequence 1, Application US/08700651B
; Patent No. 6015882
; GENERAL INFORMATION:
;  APPLICANT: PETERSEN, CAROLYN
;  APPLICANT: LEECH, JAMES
;  APPLICANT: NELSON, RICHARD, C.
;  APPLICANT: GUT, JIRI
;  TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS
;  TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
;  TITLE OF INVENTION: INFECTIONS
;  FILE REFERENCE: 480.19-4(HV)
;  CURRENT APPLICATION NUMBER: US/08/700,651B
;  CURRENT FILING DATE: 1997-08-14
;  EARLIER APPLICATION NUMBER: 08/415,751
;  EARLIER FILING DATE: 1995-04-03
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 5163
;   TYPE: DNA
;   ORGANISM: Cryptosporidium parvum
US-08-700-651-1

  Query Match             2.1%;  Score 18;  DB 3;  Length 5163;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        490 CCAATTGATCTTGAAACT 507
              ||||||||||||||||||
Db       3769 CCAATTGATCTTGAAACT 3786


RESULT 7
US-08-928-361B-4
; Sequence 4, Application US/08928361B
; Patent No. 6071518
;  GENERAL INFORMATION:
;    APPLICANT:  Petersen, Carolyn
;    TITLE OF INVENTION:  PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
;    TITLE OF INVENTION:  THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
;    TITLE OF INVENTION:  FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
;    TITLE OF INVENTION:  SPECIES INFECTIONS
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PETERS, VERNY, JONES & BIKSA
;      STREET:  385 Sherman Avenue, Suite 6
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306-1840
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/928,361B
;      FILING DATE:  12-SEP-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/026,062
;      FILING DATE:  13-SEP-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verny, Hana
;      REGISTRATION NUMBER:  30,518
;      REFERENCE/DOCKET NUMBER:  480.76-1(HV)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-324-1677
;      TELEFAX:  650-324-1678
```

```
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5163 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-928-361B-4

  Query Match             2.1%;  Score 18;  DB 3;  Length 5163;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        490 CCAATTGATCTTGAAACT 507
              ||||||||||||||||||
Db       3769 CCAATTGATCTTGAAACT 3786


RESULT 8
US-08-700-651-2
; Sequence 2, Application US/08700651B
; Patent No. 6015882
; GENERAL INFORMATION:
;  APPLICANT: PETERSEN, CAROLYN
;  APPLICANT: LEECH, JAMES
;  APPLICANT: NELSON, RICHARD, C.
;  APPLICANT: GUT, JIRI
;  TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS
;  TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
;  TITLE OF INVENTION: INFECTIONS
;  FILE REFERENCE: 480.19-4(HV)
;  CURRENT APPLICATION NUMBER: US/08/700,651B
;  CURRENT FILING DATE: 1997-08-14
;  EARLIER APPLICATION NUMBER: 08/415,751
;  EARLIER FILING DATE: 1995-04-03
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 5318
;   TYPE: DNA
;   ORGANISM: Cryptosporidium parvum
US-08-700-651-2

  Query Match             2.1%;  Score 18;  DB 3;  Length 5318;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        490 CCAATTGATCTTGAAACT 507
              ||||||||||||||||||
Db       3769 CCAATTGATCTTGAAACT 3786


RESULT 9
US-08-928-361B-3
; Sequence 3, Application US/08928361B
; Patent No. 6071518
;  GENERAL INFORMATION:
;    APPLICANT:  Petersen, Carolyn
;    TITLE OF INVENTION:  PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
;    TITLE OF INVENTION:  THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
;    TITLE OF INVENTION:  FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
;    TITLE OF INVENTION:  SPECIES INFECTIONS
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PETERS, VERNY, JONES & BIKSA
;      STREET:  385 Sherman Avenue, Suite 6
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306-1840
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/928,361B
;      FILING DATE:  12-SEP-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/026,062
;      FILING DATE:  13-SEP-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verny, Hana
;      REGISTRATION NUMBER:  30,518
;      REFERENCE/DOCKET NUMBER:  480.76-1(HV)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-324-1677
;      TELEFAX:  650-324-1678
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5318 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-928-361B-3

  Query Match             2.1%;  Score 18;  DB 3;  Length 5318;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        490 CCAATTGATCTTGAAACT 507
              ||||||||||||||||||
Db       3769 CCAATTGATCTTGAAACT 3786


RESULT 10
US-09-385-982-404
; Sequence 404, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
;  APPLICANT: ENDEGE, WILSON O., ET AL.
;  TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
;  TITLE OF INVENTION: PRODUCTS:  II
;  FILE REFERENCE: CCDNA-260XX
;  CURRENT APPLICATION NUMBER: US/09/385,982
;  CURRENT FILING DATE: 1999-08-30
;  EARLIER APPLICATION NUMBER: 09/328,111
;  EARLIER FILING DATE: 1999-06-08
;  EARLIER APPLICATION NUMBER: 60/117,393
;  EARLIER FILING DATE: 1999-01-27
;  EARLIER APPLICATION NUMBER: 60/098,639
;  EARLIER FILING DATE: 1998-08-31
;  NUMBER OF SEQ ID NOS: 544
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 404
;   LENGTH: 604
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(604)
;   OTHER INFORMATION: n = A,T,C or G
US-09-385-982-404

  Query Match             2.0%;  Score 17;  DB 4;  Length 604;
  Best Local Similarity   100.0%;  Pred. No. 35;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         83 CTCAGCTAGAAGAAAGT 99
              |||||||||||||||||
Db        289 CTCAGCTAGAAGAAAGT 305


RESULT 11
```

```
US-09-385-982-173
; Sequence 173, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
;  APPLICANT: ENDEGE, WILSON O., ET AL.
;  TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
;  TITLE OF INVENTION: PRODUCTS:  II
;  FILE REFERENCE: CCDNA-260XX
;  CURRENT APPLICATION NUMBER: US/09/385,982
;  CURRENT FILING DATE: 1999-08-30
;  EARLIER APPLICATION NUMBER: 09/328,111
;  EARLIER FILING DATE: 1999-06-08
;  EARLIER APPLICATION NUMBER: 60/117,393
;  EARLIER FILING DATE: 1999-01-27
;  EARLIER APPLICATION NUMBER: 60/098,639
;  EARLIER FILING DATE: 1998-08-31
;  NUMBER OF SEQ ID NOS: 544
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 173
    LENGTH: 616
    TYPE: DNA
    ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(616)
;   OTHER INFORMATION: n = A,T,C or G
US-09-385-982-173

  Query Match             2.0%;  Score 17;  DB 4;  Length 616;
  Best Local Similarity   100.0%;  Pred. No. 35;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         83 CTCAGCTAGAAGAAAGT 99
              |||||||||||||||||
Db        286 CTCAGCTAGAAGAAAGT 302


RESULT 12
US-08-902-516-1
; Sequence 1, Application US/08902516
; Patent No. 5891432
;  GENERAL INFORMATION:
;    APPLICANT:  Soo Hoo, William
;    TITLE OF INVENTION:  MEMBRANE-BOUND CYTOKINE COMPOSITIONS
;    TITLE OF INVENTION:  COMPRISING GM-CSF AND METHODS OF MODULATING AN IMMUNE
;    TITLE OF INVENTION:  RESPONSE USING SAME
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CAMPBELL & FLORES, LLP
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92121
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/902,516
;      FILING DATE:  29-JUL-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-IM 2442
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)535-9001
;      TELEFAX:  (619)535-8949
; INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  660 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..660
US-08-902-516-1

  Query Match             2.0%;  Score 17;  DB 2;  Length 660;
  Best Local Similarity   100.0%;  Pred. No. 35;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        633 TATGATGTTCCAGATTA 649
              |||||||||||||||||
Db         70 TATGATGTTCCAGATTA 86


RESULT 13
US-09-134-001C-2199
; Sequence 2199, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2199
;   LENGTH: 1164
;   TYPE: DNA
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2199

  Query Match             2.0%;  Score 17;  DB 4;  Length 1164;
  Best Local Similarity   100.0%;  Pred. No. 35;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        845 TAGAAATCGTTGATTTA 861
              |||||||||||||||||
Db       1001 TAGAAATCGTTGATTTA 1017


RESULT 14
US-08-809-763-3
; Sequence 3, Application US/08809763
; Patent No. 5919691
;  GENERAL INFORMATION:
;    APPLICANT:  Andersen, Lene No. 5919691boe
;    APPLICANT:  Lassen, S ren Flensted
;    APPLICANT:  Kauppinen, Markus Sakari
;    APPLICANT:  Nielsen, Jack Bech
;    TITLE OF INVENTION:  An Enzyme and Enzyme Preparation
;    TITLE OF INVENTION:  With Endoglucanase Activity
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5919691o No. 5919691disk of No. 5919691th America, Inc.
;      STREET:  405 Lexington Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10174
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US/08/809,763
;      FILING DATE:  21-MAR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lambiris, Elias J
;      REGISTRATION NUMBER:  33,728
;      REFERENCE/DOCKET NUMBER:  4257.204-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1511 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Acremonium sp.
;      STRAIN:  CBS 265.95
   08-809-763-3

  Query Match             2.0%;  Score 17;  DB 2;  Length 1511;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        489 ACCAATTGATCTTGAAA 505
              |||||||||||||||||
Db       1436 ACCAATTGATCTTGAAA 1452


RESULT 15
US-08-956-253-3
; Sequence 3, Application US/08956253
; Patent No. 6071735
;  GENERAL INFORMATION:
;    APPLICANT:  Andersen, Lene No. 6071735boe
;    APPLICANT:  Lassen, S ren Flensted
;    APPLICANT:  Kauppinen, Markus Sakari
;    APPLICANT:  Nielsen, Jack Bech
;    TITLE OF INVENTION:  An Enzyme and Enzyme Preparation
;    TITLE OF INVENTION:  With Endoglucanase Activity
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6071735o No. 6071735disk of No. 6071735th America, Inc.
;      STREET:  405 Lexington Avenue
       CITY:  New York
       STATE:  NY
       COUNTRY:  USA
;      ZIP:  10174
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/956,253
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,763
;      FILING DATE:  21-MAR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lambiris, Elias J
;      REGISTRATION NUMBER:  33,728
;      REFERENCE/DOCKET NUMBER:  4257.204-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1511 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Acremonium sp.
;      STRAIN:  CBS 265.95
US-08-956-253-3

  Query Match             2.0%;  Score 17;  DB 3;  Length 1511;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        489 ACCAATTGATCTTGAAA 505
              |||||||||||||||||
Db       1436 ACCAATTGATCTTGAAA 1452


Search completed: May 30, 2003, 19:09:07
Job time : 65 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 30, 2003, 19:07:56 ; Search time 186 Seconds
(without alignments)
6241.626 Million cell updates/sec

Title: US-09-925-637-63
Perfect score: 861
Sequence: 1 atgacgaatctaccgatgaa.....tattagaaatcgttgattta 861

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 845702 seqs, 674182571 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1691404

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		861	100.0	861	9	US-10-084-205-63	Sequence 63, Appl
2		861	100.0	861	10	US-09-925-637-63	Sequence 63, Appl
3	c	861	100.0	1893	7	US-08-781-986A-155	Sequence 155, App
4		489	56.8	882	10	US-09-815-242-8485	Sequence 8485, Ap
5		474	55.1	864	10	US-09-815-242-4184	Sequence 4184, Ap
6	c	413	48.0	413	10	US-09-815-242-2822	Sequence 2822, Ap
7	c	337	39.1	337	10	US-09-815-242-3275	Sequence 3275, Ap
8	c	19	2.2	151	10	US-09-864-761-24950	Sequence 24950, A
9	c	19	2.2	600	10	US-09-864-761-8212	Sequence 8212, Ap
10		19	2.2	130427	9	US-10-175-523-87	Sequence 87, Appl
11	c	19	2.2	640681	10	US-09-790-988-1	Sequence 1, Appli
12		18	2.1	118	9	US-10-079-854-357	Sequence 357, App
13		18	2.1	118	10	US-09-764-878-357	Sequence 357, App
14	c	18	2.1	322	9	US-10-040-739-374	Sequence 374, App
15	c	18	2.1	449	10	US-09-983-965-3145	Sequence 3145, Ap
16	c	18	2.1	721	9	US-09-764-868-109	Sequence 109, App
17	c	18	2.1	1408	9	US-09-822-846-54	Sequence 54, Appl
18		18	2.1	2064	9	US-09-822-846-55	Sequence 55, Appl
19	c	18	2.1	6638	9	US-10-091-504-1645	Sequence 1645, Ap
20	c	18	2.1	6638	10	US-09-764-869-1645	Sequence 1645, Ap
21		18	2.1	10146	7	US-08-781-986A-243	Sequence 243, App
22	c	18	2.1	99014	10	US-09-880-107-3428	Sequence 3428, Ap
23		17	2.0	105	10	US-09-864-761-23287	Sequence 23287, A
24		17	2.0	189	10	US-09-783-590-5878	Sequence 5878, Ap
25		17	2.0	345	9	US-10-046-935-1794	Sequence 1794, Ap
26		17	2.0	345	9	US-09-878-178-1794	Sequence 1794, Ap
27		17	2.0	345	9	US-10-146-502-1794	Sequence 1794, Ap
28		17	2.0	394	10	US-09-880-107-139	Sequence 139, App
29	c	17	2.0	405	9	US-10-060-036-979	Sequence 979, App
30		17	2.0	472	10	US-09-864-761-6575	Sequence 6575, Ap
31		17	2.0	489	9	US-10-060-036-2589	Sequence 2589, Ap
32	c	17	2.0	498	9	US-09-918-995-23290	Sequence 23290, A
33	c	17	2.0	512	9	US-09-918-995-20190	Sequence 20190, A
34		17	2.0	523	9	US-10-060-036-1626	Sequence 1626, Ap
35	c	17	2.0	534	9	US-10-060-036-2901	Sequence 2901, Ap
36		17	2.0	540	9	US-09-736-457-1416	Sequence 1416, Ap
37		17	2.0	540	9	US-09-902-941-1416	Sequence 1416, Ap
38		17	2.0	540	9	US-09-849-626-1416	Sequence 1416, Ap
39		17	2.0	540	9	US-10-017-754-1416	Sequence 1416, Ap
40		17	2.0	566	10	US-09-864-761-13400	Sequence 13400, A
41		17	2.0	604	9	US-09-871-161-404	Sequence 404, App
42		17	2.0	616	9	US-09-871-161-173	Sequence 173, App
43	c	17	2.0	648	10	US-09-925-301-819	Sequence 819, App
44		17	2.0	660	10	US-09-847-185-1	Sequence 1, Appli
45	c	17	2.0	721	9	US-09-898-554-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1
US-10-084-205-63
; Sequence 63, Application US/10084205
; Publication No. US20030049648A1
; GENERAL INFORMATION:
; APPLICANT: Choi, Gil
; TITLE OF INVENTION: 37 Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB515P1
; CURRENT APPLICATION NUMBER: US/10/084,205
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/151,933
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 63
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-084-205-63

Query Match 100.0%; Score 861; DB 9; Length 861;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 861; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
```

```
Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||||||||||||||||||||
Db        841 TTATTAGAAATCGTTGATTTA 861
```

RESULT 2
US-09-925-637-63
; Sequence 63, Application US/09925637
; Patent No. US20020103338A1
; GENERAL INFORMATION:
; APPLICANT: Choi
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
; FILE REFERENCE: PB560
; CURRENT APPLICATION NUMBER: US/09/925,637
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/151,933
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 08/781,986
; PRIOR FILING DATE: 1997-01-03
; PRIOR APPLICATION NUMBER: US 08/956,171
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/009,861
; PRIOR FILING DATE: 1996-01-06
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-637-63

Query Match 100.0%; Score 861; DB 10; Length 861;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 861; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||||||||||||||||||||
Db        841 TTATTAGAAATCGTTGATTTA 861
```

RESULT 3
US-08-781-986A-155/c
; Sequence 155, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,986A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Bob
;      REGISTRATION NUMBER:  30,446
;      REFERENCE/DOCKET NUMBER:  PB248PP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  155:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1893 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-781-986A-155

  Query Match           100.0%;  Score 861;  DB 7;  Length 1893;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  861;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1314 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 1255

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1254 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 1195

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          1194 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 1135

           181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1134 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 1075

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1074 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 1015

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1014 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 955

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 895

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        894 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 835

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        834 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 775

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        774 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 715

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        714 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 655

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 595

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 535

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 475

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||||||||||||||||||||
Db        474 TTATTAGAAATCGTTGATTTA 454


RESULT 4
US-09-815-242-8485
; Sequence 8485, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8485
;   LENGTH: 882
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(882)
US-09-815-242-8485

  Query Match           56.8%;  Score 489;  DB 10;  Length 882;
  Best Local Similarity  99.8%;  Pred. No. 1.9e-233;
  Matches  539;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCAT 300

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
```

```
RESULT 5
US-09-815-242-4184
; Sequence 4184, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
   APPLICANT:  Carr, Grant J.
   APPLICANT:  Yamamoto, Robert T.
   APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4184
;   LENGTH: 864
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-09-815-242-4184

  Query Match             55.1%;  Score 474;  DB 10;  Length 864;
  Best Local Similarity   99.8%;  Pred. No. 5.8e-226;
  Matches  524;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy         16 ATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTA 75
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTA 60

Qy         76 ATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATC 135
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATC 120

Qy        136 CGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATG 195
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATG 180

Qy        196 AAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCA 255
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCA 240

Qy        256 GCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAG 315
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        241 GCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCATAAAGTATATGGTGAG 300

Qy        316 TGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGT 375
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGT 360

Qy        376 GATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGT 435
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGT 420

Qy        436 GGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATT 495
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATT 480

Qy        496 GATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              |||||||||||||||||||||||||||||||||||||||||||||
Db        481 GATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 525
```

```
RESULT 6
US-09-815-242-2822/c
; Sequence 2822, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
```

```
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2822
;   LENGTH: 413
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-09-815-242-2822

  Query Match             48.0%;  Score 413;  DB 10;  Length 413;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-195;
  Matches  413;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        160 TTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATG 219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 TTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATG 354

Qy        220 ATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGA 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 ATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGA 294

Qy        280 CGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGAT 339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 CGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGAT 234

Qy        340 GCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTA 399
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 GCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTA 174

Qy        400 AAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGT 459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 AAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGT 114

Qy        460 CAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATA 519
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 CAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATA 54

Qy        520 CACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCAGA 572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         53 CACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCAGA 1


RESULT 7
US-09-815-242-3275/c
; Sequence 3275, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3275
;   LENGTH: 337
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-09-815-242-3275

  Query Match             39.1%;  Score 337;  DB 10;  Length 337;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-158;
  Matches  337;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        225 TACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGG 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 TACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGG 278

Qy        285 AAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTT 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 AAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTT 218

Qy        345 ATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAAT 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 ATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAAT 158

Qy        405 AAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAAT 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 AAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAAT 98

Qy        465 GTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAA 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 GTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAA 38

Qy        525 AACAAAAACAGGAGCATTATTAACTTTTGCGGTTATG 561
              |||||||||||||||||||||||||||||||||||||
Db         37 AACAAAAACAGGAGCATTATTAACTTTTGCGGTTATG 1


RESULT 8
US-09-864-761-24950/c
; Sequence 24950, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
```

; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 24950
; LENGTH: 151
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL117667.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
; OTHER INFORMATION: EST_HUMAN HIT: AA601023.1, EVALUE 1.00e-14
; OTHER INFORMATION: NT HIT: AJ271735.1, EVALUE 2.00e-15
; OTHER INFORMATION: SWISSPROT HIT: P08548, EVALUE 1.00e-05
US-09-864-761-24950

Query Match 2.2%; Score 19; DB 10; Length 151;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 503 AAACTTTGGAAATGATACA 521
||||||||||||||||||||
Db 139 AAACTTTGGAAATGATACA 121

ULT 9
09-864-761-8212/c
; Sequence 8212, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 8212
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL117667.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
US-09-864-761-8212

Query Match 2.2%; Score 19; DB 10; Length 600;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 503 AAACTTTGGAAATGATACA 521
||||||||||||||||||||
Db 470 AAACTTTGGAAATGATACA 452

RESULT 10
US-10-175-523-87
; Sequence 87, Application US/10175523
; Publication No. US20030096264A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, Jeffrey
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Palfreyman, Michael
; APPLICANT: Rajan, Prithi
; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 3235/1J795-US3
; CURRENT APPLICATION NUMBER: US/10/175,523
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/325,150

```
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US 60/333,047
;  PRIOR FILING DATE: 2001-11-14
;  PRIOR APPLICATION NUMBER: US 60/349,936
;  PRIOR FILING DATE: 2002-01-18
;  PRIOR APPLICATION NUMBER: US 60/361,834
;  PRIOR FILING DATE: 2002-03-04
;  NUMBER OF SEQ ID NOS: 197
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 87
;   LENGTH: 130427
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-175-523-87

  Query Match             2.2%;  Score 19;  DB 9;  Length 130427;
  Best Local Similarity   100.0%;  Pred. No. 51;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        647 TTAAAGATGATTTATTAGA 665
              |||||||||||||||||||
Db     106008 TTAAAGATGATTTATTAGA 106026


RESULT 11
US-09-790-988-1/c
; Sequence 1, Application US/09790988
; Patent No. US20020127687A1
; GENERAL INFORMATION:
;  APPLICANT: SHIGENOBU, SHUJI
;  APPLICANT:  WATANABE, HIDEMI
;  APPLICANT:  HATTORI, MASAHIRA
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
;  FILE REFERENCE: 081356/0159
;  CURRENT APPLICATION NUMBER: US/09/790,988
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: JP2000-107160
;  PRIOR FILING DATE: 2000-04-07
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 640681
;   TYPE: DNA
;   ORGANISM: Buchnera sp.
US-09-790-988-1

  Query Match             2.2%;  Score 19;  DB 10;  Length 640681;
  Best Local Similarity   100.0%;  Pred. No. 60;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 AATAAATTAATAGATGAAG 37
              |||||||||||||||||||
Db     214082 AATAAATTAATAGATGAAG 214064


RESULT 12
US-10-079-854-357
; Sequence 357, Application US/10079854
; Publication No. US20030054368A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA121C1
;  CURRENT APPLICATION NUMBER: US/10/079,854
;  CURRENT FILING DATE:  2002-02-22
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 428
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 357
;   LENGTH: 118
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-079-854-357

  Query Match             2.1%;  Score 18;  DB 9;  Length 118;
  Best Local Similarity   100.0%;  Pred. No. 71;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        601 GAACATTTAGAAAGTTAT 618
              ||||||||||||||||||
Db         14 GAACATTTAGAAAGTTAT 31


RESULT 13
US-09-764-878-357
; Sequence 357, Application US/09764878
; Patent No. US20020090615A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA121
;  CURRENT APPLICATION NUMBER: US/09/764,878
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 428
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 357
;   LENGTH: 118
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-878-357

  Query Match             2.1%;  Score 18;  DB 10;  Length 118;
  Best Local Similarity   100.0%;  Pred. No. 71;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        601 GAACATTTAGAAAGTTAT 618
              ||||||||||||||||||
Db         14 GAACATTTAGAAAGTTAT 31


RESULT 14
US-10-040-739-374/c
; Sequence 374, Application US/10040739
; Patent No. US20020173635A1
;   GENERAL INFORMATION:
;        APPLICANT: Jacobs, Kenneth
;                   McCoy, John
;                   LaVallie, Edward
;                   Racie, Lisa
;                   Merberg, David
;                   Treacy, Maurice
;                   Spaulding, Vikki
;        TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS
;        NUMBER OF SEQUENCES: 1519
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genetics Institute, Inc.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: U.S.A
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy Disk
;             COMPUTER: IBM PC Compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/040,739
;             FILING DATE: 07-Jan-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/036,520
;             FILING DATE: 03-JUN-1998
;        ATTORNEY/AGENT INFORMATION:
```

```
;             NAME: Brown, Scott A.
;             REGISTRATION NUMBER: 32,724
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617) 498-8224
;             TELEFAX: (617) 876-5851
;    INFORMATION FOR SEQ ID NO: 374:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 322 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 374:
US-10-040-739-374

  Query Match             2.1%;  Score 18;  DB 9;  Length 322;
  Best Local Similarity   100.0%;  Pred. No. 80;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

          645 GATTAAAGATGATTTATT 662
              ||||||||||||||||||
          271 GATTAAAGATGATTTATT 254


RESULT 15
US-09-983-965-3145/c
; Sequence 3145, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 37-21(10297)C
;  CURRENT APPLICATION NUMBER: US/09/983,965
;  CURRENT FILING DATE:  2001-10-26
;  PRIOR APPLICATION NUMBER: US 09/465,231
;  PRIOR FILING DATE: 1999-12-15
;  PRIOR APPLICATION NUMBER: US 60/113,678
;  PRIOR FILING DATE: 1998-12-17
;  NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 3145
;   LENGTH: 449
;   TYPE: DNA
;   ORGANISM: Bos taurus
    FEATURE:
    OTHER INFORMATION: Clone ID: 26-LIB3058-004-Q1-K1-G5
   09-983-965-3145

  Query Match             2.1%;  Score 18;  DB 10;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 83;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        291 AACAAATCATAAAGTATA 308
              ||||||||||||||||||
Db        164 AACAAATCATAAAGTATA 147


Search completed: May 30, 2003, 20:32:14
Job time : 188 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 30, 2003, 18:18:06 ; Search time 1599 Seconds
(without alignments)
8720.647 Million cell updates/sec

Title: US-09-925-637-63
Perfect score: 861
Sequence: 1 atgacgaatctaccgatgaa.....tattagaaatcgttgattta 861

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		24	2.8	317	10	BB554288	BB554288 BB554288
c 2		22	2.6	415	17	AZ481013	AZ481013 1M0302015
c 3		22	2.6	631	10	AV679322	AV679322 AV679322
4		22	2.6	652	17	AZ506985	AZ506985 1M0348H23
c 5		21	2.4	515	12	BG817859	BG817859 ESSU0281
c 6		21	2.4	1056	17	CNS06DRD	AL394175 T7 end of
c 7		20	2.3	24	17	AZ852748	AZ852748 2M0155G09
c 8		20	2.3	159	17	AQ355799	AQ355799 CITBI-E1-
9		20	2.3	216	10	BB323434	BB323434 BB323434
10		20	2.3	250	10	BB323699	BB323699 BB323699
11		20	2.3	301	10	BB188250	BB188250 BB188250
c 12		20	2.3	427	10	AW986250	AW986250 uf75g12.x
c 13		20	2.3	433	17	AQ380942	AQ380942 RPCI11-16
14		20	2.3	440	17	BH765020	BH765020 BMBAC353B
15		20	2.3	444	10	BB822482	BB822482 BB822482
16		20	2.3	449	10	BB828059	BB828059 BB828059
c 17		20	2.3	457	14	C87470	C87470 C87470 Mous
c 18		20	2.3	476	10	BE210383	BE210383 so43b02.y
19		20	2.3	522	17	AQ526183	AQ526183 HS_5312_B
c 20		20	2.3	527	12	BG067198	BG067198 H3051F10-
21		20	2.3	547	10	BB821485	BB821485 BB821485
22		20	2.3	571	17	AZ989932	AZ989932 2M0273O22
c 23		20	2.3	579	14	C88106	C88106 C88106 Mous
c 24		20	2.3	581	14	C85477	C85477 C85477 Mous
c 25		20	2.3	653	17	AQ352899	AQ352899 CITBI-E1-
c 26		20	2.3	695	17	BH664215	BH664215 BOHXA08TF
c 27		20	2.3	704	17	BH546103	BH546103 BOGFK13TF
c 28		20	2.3	734	14	BQ031102	BQ031102 UI-1-CF0-
c 29		20	2.3	748	17	AZ175417	AZ175417 SP_0132_A
30		20	2.3	840	17	BH573787	BH573787 BOGSB66TF
c 31		20	2.3	850	17	AZ528307	AZ528307 ENTCC54TR
32		20	2.3	1079	17	CNS05E6G	AL333313 Tetraodon
33		20	2.3	1362	12	BG335622	BG335622 602404245
c 34		19	2.2	114	14	BQ979812	BQ979812 QHI9E18.y
c 35		19	2.2	175	17	AL765391	AL765391 Arabidops
36		19	2.2	187	10	AW890057	AW890057 MR0-NT003
c 37		19	2.2	203	14	BQ977827	BQ977827 QHI3C06.y
c 38		19	2.2	217	14	BQ978811	BQ978811 QHI6F13.y
c 39		19	2.2	217	14	BQ979209	BQ979209 QHI7I03.y
c 40		19	2.2	222	14	BQ973041	BQ973041 QHI11L01.
c 41		19	2.2	229	14	BU034345	BU034345 QHJ3N01.y
c 42		19	2.2	230	14	BQ979463	BQ979463 QHI8D23.y
c 43		19	2.2	232	14	BQ972776	BQ972776 QHI10G06.
c 44		19	2.2	233	10	AW430406	AW430406 69917 MAR
c 45		19	2.2	244	9	AV246789	AV246789 AV246789

ALIGNMENTS

RESULT 1
BB554288
LOCUS BB554288 317 bp mRNA linear EST 01-AUG-2000
DEFINITION BB554288 RIKEN full-length enriched, 2 days pregnant adult female ovary Mus musculus cDNA clone E330012K12 3' similar to J04596 Mouse platelet-derived growth factor-inducible KC protein mRNA, mRNA sequence.
ACCESSION BB554288
VERSION BB554288.1 GI:9640654
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 317)
AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Konno,H., et al.)
JOURNAL Unpublished (2000)

COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

FEATURES Location/Qualifiers
source 1. .317
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="E330012K12"
/clone_lib="RIKEN full-length enriched, 2 days pregnant adult female ovary"
/sex="female"
/tissue_type="ovary"
/dev_stage="2 days pregnant adult"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGCGGCCGCAACTCGAGTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 78 a 56 c 68 g 114 t 1 others
ORIGIN

Query Match 2.8%; Score 24; DB 10; Length 317;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 392 ATGAAGTAAAAATAAAAGTTCTAC 415
||||||||||||||||||||||||
Db 293 ATGAAGTAAAAATAAAAGTTCTAC 316

RESULT 2
AZ481013/c
LOCUS AZ481013 415 bp DNA linear GSS 04-OCT-2000
DEFINITION 1M0302015R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0302015 R, DNA sequence.
ACCESSION AZ481013
VERSION AZ481013.1 GI:10641994
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 415)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0302 row: O column: 15
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 415.

FEATURES Location/Qualifiers
source 1. .415
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0302O15"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 151 a 62 c 80 g 122 t
ORIGIN

Query Match 2.6%; Score 22; DB 17; Length 415;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 649 AAAGATGATTTATTAGACTGCT 670
||||||||||||||||||||||
Db 32 AAAGATGATTTATTAGACTGCT 11

RESULT 3
AV679322/c
LOCUS AV679322 631 bp mRNA linear EST 05-OCT-2000
DEFINITION AV679322 Nori Satoh unpublished cDNA library Ciona intestinalis cDNA clone rcitb15k5 3', mRNA sequence.
ACCESSION AV679322
VERSION AV679322.1 GI:10117321
KEYWORDS EST.
SOURCE Ciona intestinalis.
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;

```
            Phlebobranchia; Cionidae; Ciona.
REFERENCE   1  (bases 1 to 631)
  AUTHORS   Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
  TITLE     Expressed genes in Ciona intestinalis
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Nori Satoh
            Department of Zoology
            Kyoto University
            Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
            Tel: 81-75-753-4081
            Fax: 81-75-705-1113
            Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .631
                     /organism="Ciona intestinalis"
                     /db_xref="taxon:7719"
                     /clone="rcitb15k5"
                     /clone_lib="Nori Satoh unpublished cDNA library"
                     /tissue_type="whole animal"
                     /dev_stage="tailbud"
                     /note="Vector: pBluescript SK"
BASE COUNT      204 a    119 c    125 g    183 t
ORIGIN

  Query Match             2.6%;  Score 22;  DB 10;  Length 631;
  Best Local Similarity   100.0%;  Pred. No. 7.2;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        598 AAAGAACATTTAGAAAGTTATA 619
              ||||||||||||||||||||||
Db        171 AAAGAACATTTAGAAAGTTATA 150


RESULT 4
AZ506985
LOCUS       AZ506985                652 bp    DNA     linear   GSS 05-OCT-2000
DEFINITION  1M0348H23F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0348H23 F, DNA sequence.
ACCESSION   AZ506985
VERSION     AZ506985.1  GI:10688301
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 652)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0348  row: H  column: 23
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 652.
FEATURES             Location/Qualifiers
     source          1. .652
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0348H23"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT      204 a    143 c    128 g    177 t
ORIGIN

  Query Match             2.6%;  Score 22;  DB 17;  Length 652;
  Best Local Similarity   100.0%;  Pred. No. 7.3;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        649 AAAGATGATTTATTAGACTGCT 670
              ||||||||||||||||||||||
Db         54 AAAGATGATTTATTAGACTGCT 75


RESULT 5
BG817859/c
LOCUS       BG817859                515 bp    mRNA    linear   EST 22-MAY-2001
DEFINITION  ESSU0281 S.scabiei cDNA library Sarcoptes scabiei cDNA clone
            SAS0368 5', mRNA sequence.
ACCESSION   BG817859
VERSION     BG817859.1  GI:14188839
KEYWORDS    EST.
SOURCE      Sarcoptes scabiei.
  ORGANISM  Sarcoptes scabiei
            Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
            Acariformes; Sarcoptiformes; Astigmata; Sarcoptoidea; Sarcoptidae;
            Sarcoptes.
REFERENCE   1  (bases 1 to 515)
  AUTHORS   Ljunggren,E.L., Nilsson,D., Naslund,K. and Mattsson,J.G.
  TITLE     Expressed sequence tag analysis of the parasitic mite Sarcoptes
            scabiei
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Mattsson J.G.
            Department of Parasitology (SWEPAR)
            National Veterinary Institute
            SE-751 89 Uppsala, Sweden
            Tel: +46 18 674120
            Fax: +46 18
            Email: jens.mattsson@sva.se
            Seq primer: T3 primer
            High quality sequence stop: 515.
FEATURES             Location/Qualifiers
     source          1. .515
                     /organism="Sarcoptes scabiei"
                     /db_xref="taxon:52283"
                     /clone="SAS0368"
                     /clone_lib="S.scabiei cDNA library"
                     /note="The Sarcoptes scabiei mixed lifestage library was
                     constructed by Jens G Mattsson. cDNAs were synthesized
                     from poly(A)+ RNA by oligo d(T) priming, size-selected and
                     directionally cloned into the Uni-ZAP lambda vector
                     (Stratagene). The primary library was amplified on
                     XL1-Blue MRF' cells."
BASE COUNT      199 a     72 c     92 g    152 t
```

```
ORIGIN

  Query Match             2.4%;  Score 21;  DB 12;  Length 515;
  Best Local Similarity   100.0%;  Pred. No. 21;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        814 ATTGATGAACAATTCAATACA 834
              |||||||||||||||||||||
Db        483 ATTGATGAACAATTCAATACA 463


RESULT 6
CNS06DRD/c
LOCUS       CNS06DRD                1056 bp    DNA     linear   GSS 05-APR-2001
DEFINITION  T7 end of clone AR0AA015E03 of library AR0AA from strain CBS 732 of
            Zygosaccharomyces rouxii, genomic survey sequence.
ACCESSION   AL394175
VERSION     AL394175.1  GI:12144998
KEYWORDS    GSS; clone AR0AA015E03; RST AR0AA015E03CP1.
SOURCE      Zygosaccharomyces rouxii.
  ORGANISM  Zygosaccharomyces rouxii
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Zygosaccharomyces.
REFERENCE   1  (bases 1 to 1056)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
  MEDLINE   20584711
   PUBMED   11152876
REFERENCE   2  (bases 1 to 1056)
  AUTHORS   de Montigny,J., Straub,M., Potier,S., Tekaia,F., Dujon,B.,
            Wincker,P., Artiguenave,F. and Souciet,J.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 8.
            Zygosaccharomyces rouxii
  JOURNAL   FEBS Lett. 487 (1), 52-55 (2000)
  MEDLINE   20584718
   PUBMED   11152883
REFERENCE   3  (bases 1 to 1056)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-MAR-2001) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     This BAC end sequence is part of a random genomic sequencing
            program of thirteen yeast species:
            . Saccharomyces bayanus var. uvarum,
            . Saccharomyces exiguus,
            . Saccharomyces servazzii,
            . Zygosaccharomyces rouxii,
            . Saccharomyces kluyveri,
            . Kluyveromyces thermotolerans,
            . Kluyveromyces lactis var. lactis,
            . Kluyveromyces marxianus var. marxianus
            . Pichia angusta,
            . Debaryomyces hansenii var. hansenii,
            . Pichia sorbitophila,
            . Candida tropicalis and
            . Yarrowia lipolytica.
            Genomic inserts of 3 to 5 kb were prepared and both extremities
            were sequenced. See keywords for description of this sequence and
            for the sequence of the other extremity of this insert.
FEATURES             Location/Qualifiers
     source          1. .1056
                     /organism="Zygosaccharomyces rouxii"
                     /strain="CBS 732"
                     /db_xref="taxon:4956"
                     /clone="AR0AA015E03"
                     /clone_lib="AR0AA"
                     /note="end : T7"
BASE COUNT      298 a    202 c    209 g    340 t      7 others
ORIGIN

  Query Match             2.4%;  Score 21;  DB 17;  Length 1056;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        817 GATGAACAATTCAATACAAAA 837
              |||||||||||||||||||||
Db        433 GATGAACAATTCAATACAAAA 413


RESULT 7
AZ852748/c
LOCUS       AZ852748                  24 bp    DNA     linear   GSS 21-FEB-2001
DEFINITION  2M0155G09R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0155G09 R, DNA sequence.
ACCESSION   AZ852748
VERSION     AZ852748.1  GI:13040175
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0155  row: G  column: 09
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 24.
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0155G09"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
```

chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT 5 a 7 c 3 g 9 t
ORIGIN

Query Match 2.3%; Score 20; DB 17; Length 24;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 423 GTCAATAGCAAGTGGTCATG 442
||||||||||||||||||||
Db 22 GTCAATAGCAAGTGGTCATG 3

RESULT 8
AQ355799/c
LOCUS AQ355799 159 bp DNA linear GSS 24-JAN-1999
DEFINITION CITBI-E1-2531D2.TF CITBI-E1 Homo sapiens genomic clone 2531D2, DNA sequence.
ACCESSION AQ355799
VERSION AQ355799.1 GI:4182972
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 159)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and Venter,J.C.
TITLE Use of BAC End Sequences from CalTech Libraries for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: CITBI-E1-2531D2.TR
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .159
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2531D2"
/clone_lib="CITBI-E1"
/sex="male"
/cell_type="sperm"
/note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI; CalTech Human BAC Library D"
BASE COUNT 61 a 23 c 17 g 58 t
ORIGIN

Query Match 2.3%; Score 20; DB 17; Length 159;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 TGAATAAATTAATAGATGAA 36
||||||||||||||||||||
Db 154 TGAATAAATTAATAGATGAA 135

RESULT 9
BB323434
LOCUS BB323434 216 bp mRNA linear EST 11-JUL-2000
DEFINITION BB323434 RIKEN full-length enriched, 4 days neonate male adipose Mus musculus cDNA clone B430005G12 3' similar to J04596 Mouse platelet-derived growth factor-inducible KC protein mRNA, mRNA sequence.
ACCESSION BB323434
VERSION BB323434.1 GI:9031748
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 216)
AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Konno,H., et al.)
JOURNAL Unpublished (2000)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.
FEATURES Location/Qualifiers
source 1. .216
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="B430005G12"
/clone_lib="RIKEN full-length enriched, 4 days neonate male adipose"
/sex="male"
/tissue_type="adipose"
/dev_stage="4 days neonate"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 229.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a

modified pBluescript KS(+) after bulk excision from Lambda FLC I."
BASE COUNT 53 a 45 c 36 g 82 t
ORIGIN

Query Match 2.3%; Score 20; DB 10; Length 216;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 392 ATGAAGTAAAAATAAAAGTT 411
||||||||||||||||||||
Db 191 ATGAAGTAAAAATAAAAGTT 210

RESULT 10
BB323699
LOCUS BB323699 250 bp mRNA linear EST 11-JUL-2000
DEFINITION BB323699 RIKEN full-length enriched, 4 days neonate male adipose Mus musculus cDNA clone B430011A13 3' similar to J04596 Mouse platelet-derived growth factor-inducible KC protein mRNA, mRNA sequence.
ACCESSION BB323699
VERSION BB323699.1 GI:9032013
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 250)
AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci ,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M. , Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y. , Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata ,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y. , Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya ,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I. , Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino ,M., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Konno,H., et al.)
JOURNAL Unpublished (2000)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki ,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki ,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.
FEATURES Location/Qualifiers
source 1. .250
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="B430011A13"
/clone_lib="RIKEN full-length enriched, 4 days neonate male adipose"
/sex="male"
/tissue_type="adipose"
/dev_stage="4 days neonate"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 229.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."
BASE COUNT 68 a 49 c 45 g 88 t
ORIGIN

Query Match 2.3%; Score 20; DB 10; Length 250;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 392 ATGAAGTAAAAATAAAAGTT 411
||||||||||||||||||||
Db 226 ATGAAGTAAAAATAAAAGTT 245

RESULT 11
BB188250
LOCUS BB188250 301 bp mRNA linear EST 30-JUN-2000
DEFINITION BB188250 RIKEN full-length enriched, adult male spinal cord Mus musculus cDNA clone A330047P06 3' similar to J04596 Mouse platelet-derived growth factor-inducible KC protein mRNA, mRNA sequence.
ACCESSION BB188250
VERSION BB188250.1 GI:8848821
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 301)
AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci ,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M. , Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y. , Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata ,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y. , Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya ,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I. , Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino ,M., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Konno,H., et al.)
JOURNAL Unpublished (2000)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki

,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki ,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

```
FEATURES             Location/Qualifiers
     source          1. .301
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="A330047P06"
                     /clone_lib="RIKEN full-length enriched, adult male spinal
                     cord"
                     /sex="male"
                     /tissue_type="spinal cord"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 10.0 and subtraction to Rot = 459.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC
                     3']. cDNA was cleaved with XhoI and BamHI. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I. "
BASE COUNT       73 a     55 c     68 g    105 t
ORIGIN

  Query Match             2.3%;  Score 20;  DB 10;  Length 301;
  Best Local Similarity   100.0%;  Pred. No. 57;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

        392 ATGAAGTAAAAATAAAAGTT 411
            ||||||||||||||||||||
Db      277 ATGAAGTAAAAATAAAAGTT 296


RESULT 12
AW986250/c
LOCUS       AW986250                427 bp    mRNA    linear   EST 02-JUN-2000
DEFINITION  uf75g12.x1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
            IMAGE:1528006 3', mRNA sequence.
ACCESSION   AW986250
VERSION     AW986250.1  GI:8180846
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 427)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Other_ESTs: uf75g12.y1
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:944106
            High quality sequence stop: 363.
FEATURES             Location/Qualifiers
     source          1. .427
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1528006"
                     /clone_lib="Soares_mammary_gland_NMLMG"
                     /sex="female (lactating)"
                     /tissue_type="mammary gland"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from mammary
                     gland tissue from a lactating female, and was then primed
                     with a Not I - oligo(dT) primer. Double-stranded cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not
                     I and cloned into the Not I and Eco RI sites of the
                     modified pT7T3 vector. Library is normalized. Library
                     was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      139 a     65 c     58 g    165 t
ORIGIN

  Query Match             2.3%;  Score 20;  DB 10;  Length 427;
  Best Local Similarity   100.0%;  Pred. No. 64;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          98 GTATGTTGTATTCATTAAAT 117
               ||||||||||||||||||||
Db          60 GTATGTTGTATTCATTAAAT 41


RESULT 13
AQ380942/c
LOCUS       AQ380942                433 bp    DNA     linear   GSS 21-MAY-1999
DEFINITION  RPCI11-166B24.TV RPCI-11 Homo sapiens genomic clone RPCI-11-166B24,
            DNA sequence.
ACCESSION   AQ380942
VERSION     AQ380942.1  GI:4351965
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 433)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
            ,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .433
                     /organism="Homo sapiens"
                     /db_xref="GDB:7563407"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-166B24"
```

/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"
BASE COUNT 113 a 57 c 76 g 187 t
ORIGIN

Query Match 2.3%; Score 20; DB 17; Length 433;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 502 GAAACTTTGGAAATGATACA 521
||||||||||||||||||||
Db 214 GAAACTTTGGAAATGATACA 195

RESULT 14
BH765020
US BH765020 440 bp DNA linear GSS 20-MAR-2002
INITION BMBAC353B09T7_PSU Brugia malayi Genomic Bac Library 3 Brugia malayi genomic, DNA sequence.
ACCESSION BH765020
VERSION BH765020.1 GI:19562784
KEYWORDS GSS.
SOURCE Brugia malayi.
ORGANISM Brugia malayi
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea; Onchocercidae; Brugia.
REFERENCE 1 (bases 1 to 440)
AUTHORS Whitton,C., Daub,J., Ware,J., Quail,M., Hall,N., Barrell,B., Foster ,J., Guiliano,D., Slatko,B. and Blaxter,M.
TITLE Genome survey sequences from the human parasitic nematode Brugia malayi
JOURNAL Unpublished (2000)
COMMENT Contact: Blaxter ML
Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9 3JT, UK
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk
Sequenced from the Brugia malayi BAC library constructed by Claire Whitton and Dr Mike Quail. The sequence was generated by The Pathogen Sequencing Unit, The Sanger Institute, Cambridge, UK in collaboration with Mark Blaxter, ICAPB, University of Edinburgh, Edinburgh, UK.
Seq primer: T7 (TAATACGACTCACTATAGGG)
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .440
/organism="Brugia malayi"
/strain="TRS"
/db_xref="taxon:6279"
/clone_lib="Brugia malayi Genomic Bac Library 3"
/sex="Mixed (male and female)"
/tissue_type="whole parasite"
/dev_stage="microfilaria (L1)"
/note="Vector: pBACe3.6; Site_1: BamH I; Brugia malayi genomic DNA was partially cleaved with Sau3A I and size fractionated. 7,392 clones were generated with mean insert size ~48 kbp. The library was constructed by Claire Whitton, Blaxter Nematode Genetics Lab, University of Edinburgh, UK, and Dr Mike Quail, The Pathogen Sequencing Unit, The Sanger Centre, Cambridge, UK."
BASE COUNT 151 a 55 c 56 g 178 t
ORIGIN

Query Match 2.3%; Score 20; DB 17; Length 440;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 359 TTGAACTTATTTCAAGTGAT 378
||||||||||||||||||||
Db 174 TTGAACTTATTTCAAGTGAT 193

RESULT 15
BB822482
LOCUS BB822482 444 bp mRNA linear EST 19-NOV-2001
DEFINITION BB822482 RIKEN full-length enriched, mammary gland RCB-0526 Jyg-MC(A) cDNA Mus musculus cDNA clone G830020C06 3', mRNA sequence.
ACCESSION BB822482
VERSION BB822482.1 GI:16995111
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 444)
AUTHORS Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii ,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T. , Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa ,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T. , Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. . 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. . 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara ,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
e mouse tissues.
FEATURES Location/Qualifiers
source 1. .444
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="G830020C06"
/clone_lib="RIKEN full-length enriched, mammary gland RCB-0526 Jyg-MC(A) cDNA"
/tissue_type="mammary gland"
/cell_line="RCB-0526 Jyg-MC(A)"
/note="pooled cell lines ; (cell_line=CRL-1751 WEHI 164), (cell_line=CRL-2116 JC), (cell_line=RCB-0035 WEHI-3), (cell_line=RCB-0464 Meth-A), (cell_line=RCB-0545 OHTA), (cell_line=RCB-0559 K-1 . F1), (cell_line=RCB-1283 B16 melanoma), (cell_type=B cells, cell_line=CRL-1702 WEHI 231

), (cell_type=Leydig cells, cell_line=CRL-2065 MLTC-1),
(cell_type=Nullipotent stem cell, cell_line=CRL-2070 NE),
(tissue_type=bladder, cell_line=RCB-0544 MBT-2),
(tissue_type=bone marrow, cell_type=stroma cell,
cell_line= CRL-2028 SR-4987), (tissue_type=colon,
cell_line=RCB-0549 Cle-H3), (tissue_type=kidney,
cell_line=CCL-142 RAG), (tissue_type=submandibular gland,
cell_line=CRL-1734 SCA-9 clone 15), (strain=BALB/C,
cell_type=B cells, cell_line=CRL-1669 BCL1 Clone 13.20-3B3
), (strain=C3H, tissue_type=brain, cell_line=CRL-1443
BC3H1)"

BASE COUNT 117 a 65 c 106 g 156 t
ORIGIN

Query Match 2.3%; Score 20; DB 10; Length 444;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        392 ATGAAGTAAAAATAAAAGTT 411
              ||||||||||||||||||||
          421 ATGAAGTAAAAATAAAAGTT 440
```

Search completed: May 30, 2003, 19:36:01
Job time : 1604 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 30, 2003, 16:56:55 ; Search time 185 Seconds
(without alignments)
6275.364 Million cell updates/sec

Title: US-09-925-637-63
Perfect score: 861
Sequence: 1 atgacgaatctaccgatgaa.....tattagaaatcgttgattta 861

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 845702 seqs, 674182571 residues

Total number of hits satisfying chosen parameters: 1691404

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	861	100.0	861	9	US-10-084-205-63	Sequence 63, Appl
	2	861	100.0	861	10	US-09-925-637-63	Sequence 63, Appl
c	3	861	100.0	1893	7	US-08-781-986A-155	Sequence 155, App
	4	846.6	98.3	882	10	US-09-815-242-8485	Sequence 8485, Ap
	5	831.6	96.6	864	10	US-09-815-242-4184	Sequence 4184, Ap
c	6	413	48.0	413	10	US-09-815-242-2822	Sequence 2822, Ap
c	7	337	39.1	337	10	US-09-815-242-3275	Sequence 3275, Ap
	8	196.8	22.9	882	10	US-09-815-242-6533	Sequence 6533, Ap
	9	195	22.6	7528	10	US-09-070-927A-55	Sequence 55, Appl
	10	164	19.0	811	10	US-09-974-300-1015	Sequence 1015, Ap
	11	131.4	15.3	888	10	US-09-815-242-7142	Sequence 7142, Ap
	12	124.8	14.5	876	10	US-09-815-242-9175	Sequence 9175, Ap
	13	124.2	14.4	876	10	US-09-815-242-9499	Sequence 9499, Ap
c	14	114.2	13.3	640681	10	US-09-790-988-1	Sequence 1, Appli
	15	102.4	11.9	900	10	US-09-815-242-5972	Sequence 5972, Ap
	16	98.6	11.5	891	9	US-09-941-947A-19	Sequence 19, Appl
	17	98.6	11.5	891	10	US-09-934-903-13	Sequence 13, Appl
	18	98.6	11.5	891	10	US-09-934-868-71	Sequence 71, Appl
	19	97.8	11.4	900	10	US-09-815-242-9986	Sequence 9986, Ap
	20	89.6	10.4	936	10	US-09-815-242-4260	Sequence 4260, Ap
	21	89.6	10.4	969	10	US-09-815-242-8505	Sequence 8505, Ap
	22	88.6	10.3	819	9	US-10-084-205-65	Sequence 65, Appl
	23	88.6	10.3	819	10	US-09-925-637-65	Sequence 65, Appl
	24	88.6	10.3	884	7	US-08-781-986A-185	Sequence 185, App
c	25	79.6	9.2	498	10	US-09-070-927A-509	Sequence 509, App
	26	77.4	9.0	498	10	US-09-974-300-5453	Sequence 5453, Ap
	27	76	8.8	498	10	US-09-974-300-5438	Sequence 5438, Ap
	28	75.2	8.7	1268	9	US-10-108-915-17	Sequence 17, Appl
c	29	73.4	8.5	2000	10	US-09-887-576-115	Sequence 115, App
	30	63.8	7.4	1470	9	US-10-108-915-25	Sequence 25, Appl
	31	60.8	7.1	912	10	US-09-815-242-7285	Sequence 7285, Ap
	32	60.4	7.0	1062	9	US-10-108-915-15	Sequence 15, Appl
	33	60	7.0	1161	9	US-10-108-915-37	Sequence 37, Appl
	34	57.4	6.7	1441	9	US-10-108-915-21	Sequence 21, Appl
	35	56.6	6.6	1131	10	US-09-934-778-1	Sequence 1, Appli
	36	55.6	6.5	912	9	US-09-941-947A-25	Sequence 25, Appl
	37	50.8	5.9	993	9	US-10-108-915-23	Sequence 23, Appl
	38	49.6	5.8	912	10	US-09-815-242-7450	Sequence 7450, Ap
	39	48.4	5.6	888	10	US-09-815-242-7873	Sequence 7873, Ap
	40	47.8	5.6	623	9	US-10-108-915-19	Sequence 19, Appl
	41	47.6	5.5	927	10	US-09-974-300-999	Sequence 999, App
	42	47.6	5.5	972	10	US-09-815-242-6224	Sequence 6224, Ap
	43	47.6	5.5	973	9	US-10-108-915-13	Sequence 13, Appl
	44	44.6	5.2	474	10	US-09-974-300-5445	Sequence 5445, Ap
c	45	43.6	5.1	885	10	US-09-070-927A-374	Sequence 374, App

ALIGNMENTS

RESULT 1
US-10-084-205-63
; Sequence 63, Application US/10084205
; Publication No. US20030049648A1
; GENERAL INFORMATION:
; APPLICANT: Choi, Gil
; TITLE OF INVENTION: 37 Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB515P1
; CURRENT APPLICATION NUMBER: US/10/084,205
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/151,933
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 63
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-084-205-63

Query Match 100.0%; Score 861; DB 9; Length 861;
Best Local Similarity 100.0%; Pred. No. 7.8e-175;
Matches 861; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      •181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
```

```
Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||||||||||||||||||||
Db        841 TTATTAGAAATCGTTGATTTA 861
```

RESULT 2
US-09-925-637-63
; Sequence 63, Application US/09925637
; Patent No. US20020103338A1
; GENERAL INFORMATION:
; APPLICANT: Choi
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
; FILE REFERENCE: PB560
; CURRENT APPLICATION NUMBER: US/09/925,637
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/151,933
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 08/781,986
; PRIOR FILING DATE: 1997-01-03
; PRIOR APPLICATION NUMBER: US 08/956,171
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/009,861
; PRIOR FILING DATE: 1996-01-06
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-637-63

Query Match 100.0%; Score 861; DB 10; Length 861;
Best Local Similarity 100.0%; Pred. No. 7.8e-175;
Matches 861; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||||||||||||||||||||
Db        841 TTATTAGAAATCGTTGATTTA 861
```

RESULT 3
US-08-781-986A-155/c
; Sequence 155, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,986A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Bob
;      REGISTRATION NUMBER:  30,446
;      REFERENCE/DOCKET NUMBER:  PB248PP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  155:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1893 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-781-986A-155

  Query Match             100.0%;  Score 861;  DB 7;  Length 1893;
  Best Local Similarity   100.0%;  Pred. No. 1e-174;
  Matches  861;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1314 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 1255

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1254 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 1195

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          1194 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 1135

           181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1134 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 1075

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1074 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 1015

Qy        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1014 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 955

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 895

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        894 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 835

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        834 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 775

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        774 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 715

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        714 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 655

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 595

Qy        721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 535

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 475

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||||||||||||||||||||
Db        474 TTATTAGAAATCGTTGATTTA 454


RESULT 4
US-09-815-242-8485
; Sequence 8485, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8485
;   LENGTH: 882
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(882)
US-09-815-242-8485

  Query Match             98.3%;  Score 846.6;  DB 10;  Length 882;
  Best Local Similarity   99.0%;  Pred. No. 9.4e-172;
  Matches  852;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;
```

```
Qy          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGACGAATCTACCGATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCG 60

Qy         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAAATAAATCAGTAATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCT 120

Qy        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAG 180

Qy        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATT 240

Qy        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCAT 300
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        241 CATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCAT 300

          301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTT 360

Qy        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGG 420

Qy        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGC 480

Qy        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCA 540

Qy        541 TTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAA 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTATTAACATTTGCGGTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAA 600

Qy        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAACATTTAGAAAGTTATAGTTATCATTTAGGTATGATGTTTCAGATTAAAGATGATTTA 660

Qy        661 TTAGACTGCTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTAGACTGCTATGGTGATGAAGCGAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAAT 720

          721 AATAAAAGTACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTAT 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATAAAAGTACATACGTGAGTTTATTAGGAAAAGATGGCGCAGAAGATAAATTGACTTAT 780

Qy        781 CATAGAGACGCAGCAGTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATAGAGACGCAGCAGTGGATGAACTAATGCAAATTGATGAACAATTCAATACAAAACAC 840

Qy        841 TTATTAGAAATCGTTGATTTA 861
              |||||||||||||||||||||
Db        841 TTATTAGAAATCGTTGATTTA 861
```

```
RESULT 5
US-09-815-242-4184
; Sequence 4184, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4184
;   LENGTH: 864
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-09-815-242-4184

  Query Match             96.6%;  Score 831.6;  DB 10;  Length 864;
  Best Local Similarity   98.9%;  Pred. No. 1.5e-168;
  Matches  837;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;
```

```
Qy         16 ATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTA 75
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAATAAATTAATAGATGAAGTCAATAATGAATTATCGGTTGCGATAAATAAATCAGTA 60

Qy         76 ATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATC 135
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATC 120

Qy        136 CGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATG 195
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATG 180

Qy        196 AAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCA 255
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCA 240

Qy        256 GCGATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAG 315
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        241 GCGATGGATAATGATGATTATCGACGAGGGAAATTAACAAATCATAAAGTATATGGTGAG 300

Qy        316 TGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGT 375
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGT 360

Qy        376 GATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGT 435
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATGATAGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGT 420

Qy        436 GGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATT 495
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATT 480

Qy        496 GATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTTTTGCG 555
              |||||||||||||||||||||||||||||||||||||||||||| ||||||| ||||||
Db        481 GATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCACTATTAACATTTGCG 540

Qy        556 GTTATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAGAAAGT 615
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        541 GTTATGAGTGCAGCAGATATCGCTAATGTCGATGATGCAACTAAAGAACATTTAGAAAGT 600

Qy        616 TATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCTATGGT 675
```

```
Db        601 TATAGTTATCATTTAGGTATGATGTTTCAGATTAAAGATGATTTATTAGACTGCTATGGT 660

Qy        676 GATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTACGTAC 735

Db        661 GATGAAGCGAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTACATAC 720

Qy        736 GTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGACGCAGCA 795

Db        721 GTGAGTTTATTAGGAAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGACGCAGCA 780

Qy        796 GTGGATGAACTAACGCAAATTGATGAACAATTCAATACAAAACACTTATTAGAAATCGTT 855

Db        781 GTGGATGAACTAATGCAAATTGATGAACAATTCAATACAAAACACTTATTAGAAATCGTT 840

Qy        856 GATTTA 861

Db        841 GATTTA 846


SULT 6
-09-815-242-2822/c
; Sequence 2822, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
   PRIOR APPLICATION NUMBER: 60/257,931
   PRIOR FILING DATE: 2000-12-22
   PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2822
;   LENGTH: 413
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-09-815-242-2822

  Query Match             48.0%;  Score 413;  DB 10;  Length 413;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-79;
  Matches  413;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        160 TTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATG 219

Db        413 TTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATG 354

Qy        220 ATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGA 279

Db        353 ATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGA 294

Qy        280 CGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGAT 339

Db        293 CGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGAT 234

Qy        340 GCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTA 399

Db        233 GCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTA 174

Qy        400 AAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGT 459

Db        173 AAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGT 114

Qy        460 CAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATA 519

Db        113 CAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATA 54

Qy        520 CACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCAGA 572

Db         53 CACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCAGA 1


RESULT 7
US-09-815-242-3275/c
; Sequence 3275, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3275
;   LENGTH: 337
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-09-815-242-3275

  Query Match             39.1%;  Score 337;  DB 10;  Length 337;
  Best Local Similarity   100.0%;  Pred. No. 6e-63;
  Matches  337;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        225 TACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGG 284

Db        337 TACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGACGAGG 278

Qy        285 AAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTT 344

Db        277 AAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGATGCTTT 218

Qy        345 ATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAAT 404
```

```
Db        217 ATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTAAAAAT 158

Qy        405 AAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAAT 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 AAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGTCAAAT 98

Qy        465 GTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAA 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 GTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATACACAA 38

Qy        525 AACAAAAACAGGAGCATTATTAACTTTTGCGGTTATG 561
              |||||||||||||||||||||||||||||||||||||
Db         37 AACAAAAACAGGAGCATTATTAACTTTTGCGGTTATG 1


RESULT 8
US-09-815-242-6533
; Sequence 6533, Application US/09815242
; Patent No. US20020061569A1
;   GENERAL INFORMATION:
;   APPLICANT: Haselbeck, Robert
;   APPLICANT:  Ohlsen, Kari L.
;   APPLICANT:  Zyskind, Judith W.
;   APPLICANT:  Wall, Daniel
;   APPLICANT:  Trawick, John D.
;   APPLICANT:  Carr, Grant J.
;   APPLICANT:  Yamamoto, Robert T.
;   APPLICANT:  Xu, H. Howard
;   TITLE OF INVENTION: Identification of Essential Genes in
;   TITLE OF INVENTION:  Prokaryotes
;   FILE REFERENCE: ELITRA.011A
;   CURRENT APPLICATION NUMBER: US/09/815,242
;   CURRENT FILING DATE:  2001-03-21
;   PRIOR APPLICATION NUMBER: 60/191,078
;   PRIOR FILING DATE: 2000-03-21
;   PRIOR APPLICATION NUMBER: 60/206,848
;   PRIOR FILING DATE: 2000-05-23
;   PRIOR APPLICATION NUMBER: 60/207,727
;   PRIOR FILING DATE: 2000-05-26
;   PRIOR APPLICATION NUMBER: 60/242,578
;   PRIOR FILING DATE: 2000-10-23
;   PRIOR APPLICATION NUMBER: 60/253,625
;   PRIOR FILING DATE: 2000-11-27
;   PRIOR APPLICATION NUMBER: 60/257,931
;   PRIOR FILING DATE: 2000-12-22
;   PRIOR APPLICATION NUMBER: 60/269,308
;   PRIOR FILING DATE: 2001-02-16
;   NUMBER OF SEQ ID NOS: 14110
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6533
;   LENGTH: 882
;   TYPE: DNA
;   ORGANISM: Enterococcus faecalis
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(882)
US-09-815-242-6533

  Query Match           22.9%;  Score 196.8;  DB 10;  Length 882;
  Best Local Similarity   56.1%;  Pred. No. 8e-33;
  Matches  413;  Conservative    0;  Mismatches  317;  Indels    6;  Gaps    2;

Qy         92 AAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTAT 151
              |||||   ||| | || ||| || | || || || |||||  | || ||  |  |  |
Db         89 AAGAAGCAATGCTCTACTCAATACACGCAGGTGGCAAACGATTACGCCCGCTATTAGTGC 148

Qy        152 TACTCACTTTAGATTCACTAAATACCGAGTATGA---GTTAGGTATGAAGAGCGCAATTG 208
              ||   ||  | |   |  |  | |  |||   ||   |  ||      |    ||
Db        149 TAACAACAGTGGCCGCTTTTCAAAAAGAGATGGAAACGCAAGACTATCAAGTGGCTGCCT 208

Qy        209 CACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATG 268
              |  |||| |||||||||||| |||||| | ||||||||||||  ||||||| |||||| |||

Db        209 CTTTAGAGATGATTCATACGTATTCATTAATTCATGATGATTTACCAGCAATGGATGATG 268

Qy        269 ATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATAT 328
              ||||||  || || || |||  ||| ||||| ||||| | ||||||   |||||| || |
Db        269 ATGATTTACGTCGTGGCAAACCAACCAATCACAAAGTGTTTGGTGAAGCGACTGCCATTT 328

Qy        329 TAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAA 388
              |||| || || |  ||||||||    ||||||  |  |  ||||   |         ||||
Db        329 TAGCTGGGGACGGCTTATTAACAGGCGCATTTCAGTTACTTTC---TTTGAGCCAATTAG 385

Qy        389 CTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAA 448
                  | |||| |   | |  | ||||  ||| |||  |||   ||  ||   || |
Db        386 GCTTAAGTGAAAAAGTTTTACTGATGCAACAACTGGCAAAGGCAGCGGGGAATCAAGGCA 445

Qy        449 TGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTT 508
              |||| | ||| ||||||   |||||  || | |||           |   | | |||   |
Db        446 TGGTTGCCGGCCAAATGGGTGATATTGAAGGAGAAAAAGTCAGCTTGACTTTAGAAGAGT 505

Qy        509 TGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTATGAGTGCAG 568
              | |       | ||  |||  ||||| |||||  || |    || ||  | ||   || |
Db        506 TAGCGGCTGTTCATGAAAAGAAAACCGGAGCGCTAATTGAATTCGCCTTAATTGCTGGGG 565

Qy        569 CAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAGAAAGTTATAGTTATCATT 628
                |   | ||  |||       ||   |  | |       |   |  | |    |||| |
Db        566 GCGTATTAGCCAATCAAACAGAAGAAGTCATTGGTTTACTGACACAGTTTGCGCATCACT 625

Qy        629 TAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCTATGGTGATGAAGCAAAGT 688
              ||  ||   || ||  |||   ||||| ||||||||   |    ||   |||||| | |
Db        626 ATGGCTTGGCTTTTCAAATTCGTGATGACTTATTAGATGCCACAAGTACAGAAGCAGATT 685

Qy        689 TAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTACGTACGTGAGTTTATTAG 748
              ||||||||||||| ||| | |||   |     ||||||||||||||      | |  |||
Db        686 TAGGTAAAAAAGTAGGCCGAGATGAAGCGTTGAATAAAAGTACGTATCCAGCTCTTCTAG 745

Qy        749 GGAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGACGCAGCAGTGGATGAACTAA 808
              |||  | ||||||  |||||    | ||  ||||   ||  | ||     | ||   ||
Db        746 GGATTGCTGGCGCGAAAGATGCGCTAACACATCAATTAGCAGAAGGCAGCGCTGTTTTAG 805

Qy        809 CGCAAATTGATGAACA 824
                |||| | | | |
Db        806 AAAAAATCAAAGCAAA 821


RESULT 9
US-09-070-927A-55
; Sequence 55, Application US/09070927A
; Patent No. US20020120116A1
;   GENERAL INFORMATION:
;        APPLICANT: Charles A. Kunsch
;                   Patrick J. Dillon
;                   Steven Barash
;        TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
;        NUMBER OF SEQUENCES: 982
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/070,927A
;             FILING DATE: 04-May-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/046,655
;             FILING DATE: 1997-05-16
;             APPLICATION NUMBER: 60/044,031
```

```
;             FILING DATE: 1997-05-06
;             APPLICATION NUMBER: 60/066,009
;             FILING DATE: 1997-11-14
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Kenley K. Hoover
;             REGISTRATION NUMBER: 40,302
;             REFERENCE/DOCKET NUMBER: PB369
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (301) 309-8504
;             TELEFAX: (301) 309-8512
;   INFORMATION FOR SEQ ID NO: 55:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 7528 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 55:
US-09-070-927A-55

  Query Match             22.6%;  Score 195;  DB 10;  Length 7528;
  Best Local Similarity   56.9%;  Pred. No. 4.3e-32;
  Matches  398;  Conservative    0;  Mismatches  295;  Indels    6;  Gaps    2;

Qy         92 AAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTAT 151
              |||||   ||| | || ||| || | || || || |||||  | || ||  |  |  | |
Db       6730 AAGAAGCAATGCTCTACTCAATACACGCAGGTGGCAAACGATTACGCCCGCTATTAGTGT 6789

Qy        152 TACTCACTTTAGATTCACTAAATACCGAGTATGA---GTTAGGTATGAAGAGCGCAATTG 208
              ||   ||  | |   |  |  | |  |||   ||      ||      |    ||
Db       6790 TAACAACAGTGGCCGCTTTTCAAAAAGAGATGGAAACACAAGACTATCAAGTGGCTGCCT 6849

Qy        209 CACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATG 268
              |  |||| |||||||||||| |||||| | ||||||||||| |||||| |||||  |||
Db       6850 CTTTAGAGATGATTCATACGTATTCATTAATTCATGATGATTTACCAGCAATGGACGATG 6909

Qy        269 ATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATAT 328
              ||||||  || || || |||  ||| |||||||||||| | ||||||   |||||| || |
Db       6910 ATGATTTACGTCGTGGCAAACCAACCAATCATAAAGTGTTTGGTGAAGCGACTGCCATTT 6969

Qy        329 TAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAA 388
              |||||||  || |  |||||||    |||||| |  |  ||||   |      |  ||||
Db       6970 TAGCAGGGGACGGCTTATTAACAGGTGCATTTCAGTTGCTTTC---TTTGAGTCAATTAG 7026

Qy        389 CTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAA 448
                  |||| |   | |  | |||| ||| |||  |||   ||  ||   || |
Db       7027 GCTTAAGTGAAAAAGTTTTACTGATGCAACAACTGGCAAAGGCAGCGGGGAATCAAGGCA 7086

Qy        449 TGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTT 508
              ||||  ||| ||||||   |||||  || | |||           |  | | |||   |
Db       7087 TGGTTTCCGGCCAAATGGGTGATATTGAAGGAGAAAAAGTCAGCTTGACTTTAGAAGAGT 7146

Qy        509 TGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTATGAGTGCAG 568
              | |      | ||  |||  ||||| |||||  || |    || ||  | ||   ||  |
Db       7147 TAGCGGCTGTTCATGAAAAGAAAACCGGAGCGCTAATTGAATTCGCCTTAATTGCTGGGG 7206

Qy        569 CAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAGAAAGTTATAGTTATCATT 628
                |   | ||  |||      ||  |  | |  |  |  |  | |    ||||  |
Db       7207 GCGTATTAGCCAATCAAACAGAAGAAGTCATTGGTTTACTGACACAGTTTGCGCATCACT 7266

Qy        629 TAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCTATGGTGATGAAGCAAAGT 688
                ||  ||   || || |||  ||||| ||||||||    |    ||   |||||| | |
Db       7267 ATGGCTTGGCTTTTCAAATTCGTGATGACTTATTAGATGCCACAAGTACAGAAGCAGATT 7326

Qy        689 TAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTACGTACGTGAGTTTATTAG 748
              |||||||||||||  ||| | |||   |     ||||||||||||||       | |  |||
Db       7327 TAGGTAAAAAAGTAGGCCGAGATGAAGCGTTGAATAAAAGTACGTATCCAGCTCTTCTAG 7386

Qy        749 GGAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAG 787
              |||  | ||||||  |||||    | ||  ||||   ||
Db       7387 GGATTGCTGGCGCGAAAGATGCGCTAACACATCAATTAG 7425

RESULT 10
```

```
US-09-974-300-1015
; Sequence 1015, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
;  APPLICANT: Berka, Randy M.
;  APPLICANT:  Clausen, Ib Groth
;  TITLE OF INVENTION: Methods For Monitoring Multiple Gene
;  TITLE OF INVENTION:  Expression
;  FILE REFERENCE: 10085.500-US
;  CURRENT APPLICATION NUMBER: US/09/974,300
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: 09/680,598
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/279,526
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 8481
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1015
;   LENGTH: 811
;   TYPE: DNA
;   ORGANISM: Bacillus licheniformis
US-09-974-300-1015

  Query Match             19.0%;  Score 164;  DB 10;  Length 811;
  Best Local Similarity   52.9%;  Pred. No. 8e-26;
  Matches  378;  Conservative    0;  Mismatches  330;  Indels    6;  Gaps    1;

Qy         88 CTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCGACCAGTTCTG 147
              ||  | |||   ||| |||| |||||  | || || || |||||  |  |||| |  ||
Db         97 CTCAAGGAATCCATGCTGTACTCATTGGAAGCCGGGGGGAAACGGCTGAGACCGATATTG 156

Qy        148 TTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATT 207
                |  |      | |  || |      |  ||  | ||    || ||      || |
Db        157 GTTCTTGCGCTTCTTCATGCGTACGGAAAAGACGAAGAAGCCGGAATTCCAGTCGGATGC 216

Qy        208 GCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAAT 267
              ||  | |||||||| ||||| |||||| | || |||||  |||||  ||    ||||||  |
Db        217 GCGGTTGAAATGATCCATACGTATTCATTAATCCATGACGACCTCCCTTGTATGGATGAC 276

Qy        268 GATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATA 327
              |||||     || |||| ||    ||||| || ||| | || || |||  ||| || ||
Db        277 GATGACCTCAGAAGAGGTAAGCCGACAAACCACAAAATTTACGGGGAGGCGACGGCCATT 336

Qy        328 TTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTC------AAGTGATGAT 381
              | || || || || ||||||||  ||   ||  || | ||  |      |   |  |
Db        337 CTCGCGGGAGACGCCTTATTAACCGAAAGCTTCAAAATGATCACTTCCAATATGCCTTCT 396

Qy        382 AGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCAT 441
                 || | | ||||      || | | |  |  |||||       |||   ||
Db        397 GATGTATCCGCTGAAAAGCGCATCAGACTGGTGAACGAGCTGATTTCGGCAGCGGGGGCC 456

Qy        442 GTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTT 501
              |  || |||||||||||||| || |||||||| ||   |||    |  |  |  ||||
Db        457 GAAGGCATGGTCGGCGGTCAGATTTTAGATATGGAAGCGGAATCAAAGTCTGTATCTCTT 516

Qy        502 GAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTATG 561
              ||      || |    || ||| ||  |||||| |    | |  | |  ||    || ||
Db        517 GATGAACTGCAGCGAATCCACGAAGGAAAAACGGCTAAACTCCTCAGCTTCAGCGTCATT 576

Qy        562 AGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAGAAAGTTATAGT 621
                  | |||   | | ||  |||    ||| | |    |  ||| |  |         | ||
Db        577 GCGGGAGCCATTCTTGCAGATGCATCTGAAAAAGAGATCGAAAAGCTGCGCGAATTCAGC 636

Qy        622 TATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCTATGGTGATGAA 681
              |||||  | || ||    || ||||| | ||||||  | || |||    | ||    ||
Db        637 CATCATATCGGCATCGGCTTTCAGATCAGAGATGACATTTTGGACCTTGAAGGCTCGGAG 696

Qy        682 GCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTACGTACGTGAGT 741
              |  ||| | || |||    | ||   |||   |  |||    ||    |||||   |
Db        697 GATAAGATCGGCAAACGGATCGGATCCGACGCTTCAAACGGAAAGTCGACGTATCCGTCG 756

Qy        742 TTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGACGCAGCA 795
              |  |   |  || |||||     | ||| |   | | ||||  ||   ||||
```

```
Db        757 CTCCTTTCGCTTGAGGGCGCCAGCCAAAAACTTGATGAACATATTGAAAAAGCA 810


RESULT 11
US-09-815-242-7142
; Sequence 7142, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
   CURRENT APPLICATION NUMBER: US/09/815,242
   CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7142
;   LENGTH: 888
;   TYPE: DNA
;   ORGANISM: Haemophilus influenzae
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(888)
US-09-815-242-7142

  Query Match             15.3%;  Score 131.4;  DB 10;  Length 888;
  Best Local Similarity   49.9%;  Pred. No. 7.8e-19;
  Matches  361;  Conservative    0;  Mismatches  356;  Indels    6;  Gaps    1;

Qy        100 ATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCGACCAGTTCTGTTATTACTCACT 159
              |||   ||  |||||     ||| ||||| ||  | || ||  |  |  | |     ||
Db        109 ATGAAATACGCATTATTACTTGGCGGTAAGCGAGTTCGCCCTTTCTTAGTTTACGCAACA 168

Qy        160 TTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATTGCACTAGAAATG 219
                  |    ||      | ||| |  |          |     ||   |||  | |||
Db        169 GGTCAAATGCTTGGCGCAGAGAAACAAACCTTAGATTACGCTGCTGCTGCCATTGAAGCC 228

Qy        220 ATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAATGATGATTATCGA 279
              |||||  | |||||  | ||||| |||||  |||| || |||||| ||||  |||  ||
Db        229 ATTCACGCCTATTCCTTAATTCACGATGATTTACCTGCAATGGATGATGACAATTTACGC 288

Qy        280 CGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATATTAGCAGGTGAT 339
              || ||| |    ||   |||||    || || |||    || || ||  | || || |||
Db        289 CGTGGACATCCTACTTGTCATATCCAATTTGATGAAGCTACAGCTATTCTTGCGGGCGAT 348

Qy        340 GCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGATTAACTGATGAAGTA 399
              ||  |   || |   |||||||||  |  |  | |         |  |  ||  ||||  |
Db        349 GCACTGCAAAGTTTCGCATTTGAAATATTAACCAAAACACCGAATATTTCTACTGAACAA 408

Qy        400 AAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTGGAATGGTCGGCGGT 459
              ||| |     |  | |||    |  |   ||   ||||      ||||||      ||
Db        409 AAACTGGCTTTAATTCAAATTTTAGCGCAAGGCGCTGGTGTGCAAGGAATGTGTTTAGGG 468

Qy        460 CAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAACTTTGGAAATGATA 519
              ||||  |||||| |        ||     |||  |||  | |        || ||| | ||
Db        469 CAAAGTTTAGATCTTATTTCTGAGCATAAACAGATTAGTTTAAGTGAATTAGAATTAATT 528

Qy        520 CACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTATGAGTGCAGCAGATATCGC- 578
              ||    |  ||||| || |||||  ||| |   ||  | |     |      |    .|
Db        529 CATCGTAACAAAACGGGTGCATTGCTAATTGCCGCATTGAAATTAGGTTTCATTTGTTCT 588

Qy        579 -----TAATGTCGATGATACAACTAAAGAACATTTAGAAAGTTATAGTTATCATTTAGGT 633
                  |  |  |||  | |       || ||   |||  |   |||    |      | |||
Db        589 CCGCATTTTACCGACAAAAGGTTAGAACAATCCTTAACACAATATGCGGAAGCCATTGGT 648

Qy        634 ATGATGTTCCAGATTAAAGATGATTTATTAGACTGCTATGGTGATGAAGCAAAGTTAGGT 693
                |    || ||  || |||| ||| | |||||     | || |||   ||| |  | ||
Db        649 TTAGCCTTTCAAGTTCAAGACGATATTTTAGATATTGAAGGCGATAGTGCAGAAATTGGC 708

Qy        694 AAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTACGTACGTGAGTTTATTAGGGAAA 753
              ||| | |||||   |||||||||     ||||||||||  ||    |  ||| | ||  |
Db        709 AAACAGGTGGGTGCCGATCTTGATTTAGATAAAAGTACATATCCAAAATTACTTGGATTA 768

Qy        754 GATGGCGCAGAAGATAAATTGACTTATCATAGAGACGCAGCAGTGGATGAACTAACGCAA 813
              |||. ||| || | |||  |    |||       |  ||  ||  |||| ||    ||
Db        769 AGTGGTGCAAAACAAAAAGCGCAAGATCTATATCAAAGTGCGTTGTCTGAATTAGAAAAA 828

Qy        814 ATT 816
              |||
Db        829 ATT 831


RESULT 12
US-09-815-242-9175
; Sequence 9175, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9175
;   LENGTH: 876
;   TYPE: DNA
;   ORGANISM: Streptococcus pneumoniae
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(876)
US-09-815-242-9175
```

```
  Query Match             14.5%;  Score 124.8;  DB 10;  Length 876;
  Best Local Similarity   51.2%;  Pred. No. 2e-17;
  Matches  344;  Conservative    0;  Mismatches  322;  Indels    6;  Gaps    2;

Qy         89 TAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCGACCAGTTCTGT 148
              ||   ||   | |  | |||||  |  ||||||| || || || || || ||  |||| |
Db         80 TACGGGAGTCTGTTCTCTATTCTATTCATGCTGGTGGCAAGCGTATTCGGCCTTTTCTCT 139

Qy        149 TATTACTCACTTTAGATTCACTAAA---TACCGAGTATGAGTTAGGTATGAAGAGCGCAA 205
              | |||     | | ||  |  | |   ||     |             | ||   || |
Db        140 TGTTAGAAGTTCTGGAAGCCTTGCAGGTTGCCATCAAACCTGCTCACGCGCAGGTAGCTA 199

Qy        206 TTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATA 265
               |||  | || ||||||||||||       | ||||| |||||||| || || ||||||
Db        200 CTGCCTTGGAGATGATTCATACAGGGAGCTTGATTCACGATGACCTTCCTGCTATGGATG 259

Qy        266 ATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGA 325
              |||| |||||||| ||||    ||||| ||||| ||   ||  |||||    |  || |
Db        260 ATGACGATTATCGAAGAGGGCGGTTAACCAATCACAAGAAATTCGGTGAAGCTATGGCCA 319

Qy        326 TATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGAT 385
              | || || || ||||| |||||     |  ||| ||  | ||| |    |  |||   |
Db        320 TTTTGGCTGGAGATGCCTTATTCTTAGACCCATATGCCTTGATTGCGCAGGCAGAT---T 376

Qy        386 TAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTG 445
              |  |   | |  | ||  |  | |  |        | ||  | ||    |||  | | |
Db        377 TGCCAAGTCAGATCAAGGTGGACTTGATTGCCAACTTATCCCTTGCTTCAGGTAGTCTGG 436

Qy        446 GAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAA 505
              | |||||  |  || ||| | || ||||||  |  |||||  |||  |   |   ||| |||
Db        437 GTATGGTGGCAGGGCAAGTTTTGGATATGGAGGGCGAACACCAGCACTTGTCTCTGGAAG 496

Qy        506 CTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTATGAGTG 565
                |  |  || ||    |  || || ||    ||| || | | | |  |        |
Db        497 AACTTCAGACTATTCATGCCAATAAGACTGGGAAGTTACTAGCCTATCCCTTCCAAGCGG 556

Qy        566 CAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAGAAAGTTATAGTTATC 625
              ||||   ||| ||| |  |      |  ||   | |   |  | ||| |    || |
Db        557 CAGCTATTATAGCTGAATTGTCACCTGAAATGCAGGTGAAGCTGAAAACTGTGGGTGAAT 616

Qy        626 ATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCTATGGTGATGAAGCAA 685
                | ||  |    || ||  | | ||||||| || | ||       | |  |   |
Db        617 TGATTGGACTTGCTTTTCAAGTCAGAGATGATGTACTGGATGTGACAGCTAGTTTTGAGG 676

Qy        686 AGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTACGTACGTGAGTTTAT 745
              |  | ||||| | |      | |||||  |     | |||   || ||       || |
Db        677 AAATCGGTAAGACACCTCAAAAGGATCTGCAGGCAGAAAAATCAACCTATCCTGCCTTGT 736

          746 TAGGGAAAGATG 757
              | ||    || |
          737 TGGGCTTGGAAG 748


RESULT 13
US-09-815-242-9499
; Sequence 9499, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9499
;   LENGTH: 876
;   TYPE: DNA
;   ORGANISM: Streptococcus pneumoniae
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(876)
US-09-815-242-9499

  Query Match             14.4%;  Score 124.2;  DB 10;  Length 876;
  Best Local Similarity   51.6%;  Pred. No. 2.7e-17;
  Matches  335;  Conservative    0;  Mismatches  308;  Indels    6;  Gaps    2;

Qy         89 TAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCGACCAGTTCTGT 148
              ||   ||   | |  | |||||  |  ||||||| || || || || || ||  |||| |
Db         80 TACGGGAGTCTGTTCTCTATTCTATTCATGCTGGTGGCAAGCGTATTCGGCCTTTTCTCT 139

Qy        149 TATTACTCACTTTAGATTCACTAAA---TACCGAGTATGAGTTAGGTATGAAGAGCGCAA 205
              | |||     | | ||  |  | |   ||||    |             | ||   || |
Db        140 TGTTAGAAGTTCTGGAAGCCTTGCAGGTTACCATCAAACCTGCTCACGCGCAGGTAGCTA 199

Qy        206 TTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATA 265
               |||  | || ||||||||||||       | ||||| |||||||| || || ||||||
Db        200 CTGCCTTGGAGATGATTCATACAGGGAGCTTGATTCACGATGACCTTCCTGCTATGGATG 259

Qy        266 ATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGA 325
              |||| |||||||| ||||    ||||| ||||| ||   ||  |||||    |  || |
Db        260 ATGACGATTATCGAAGAGGGCGGTTAACCAATCACAAGAAATTCGGTGAAGCTATGGCCA 319

Qy        326 TATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATGATAGAT 385
              | || || || ||||| |||||     |  ||| ||  | ||| |    |  |||   |
Db        320 TTTTGGCTGGAGATGCCTTATTCTTAGACCCATATGCCTTGATTGCGCAGGCAGAT---T 376

Qy        386 TAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCATGTTG 445
              |  |   | |  | ||  |  | |  |        | ||  | ||    |||  | | |
Db        377 TGCCAAGTCAGATTAAGGTGGACTTGATTGCCAACTTATCCCTTGCTTCAGGTAGTCTGG 436

Qy        446 GAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTTGAAA 505
              | |||||  |  || ||| | || ||||||  |  |||||  |||  |   |   | | |||
Db        437 GTATGGTGGCAGGGCAAGTTTTGGATATGGAGGGCGAACACCAGCACTTGTCTTTGGAAG 496

Qy        506 CTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTATGAGTG 565
                |  | || || ||    |  || || ||    ||| || | | | |  |        |
Db        497 AACTTCAGATTATTCATGCCAATAAGACTGGGAAGTTACTAGCCTATCCCTTCCAAGCGG 556

Qy        566 CAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTTAGAAAGTTATAGTTATC 625
              ||||   ||| ||| |  |      |  ||   | |   |  | ||| |    || |
Db        557 CAGCTATTATAGCTGAATTGTCACCTGAAATGCAGGTGAAGCTGAAAACTGTGGGTGAAT 616

Qy        626 ATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCTATGGTGATGAAGCAA 685
                | ||  |    || ||  | | ||||||| || | ||       | |  |   |
Db        617 TGATTGGACTTGCTTTTCAAGTCAGAGATGATGTACTGGATGTGACAGCTAGTTTTGAGG 676

Qy        686 AGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAGTACGTA 734
              |  | ||  || | |      |  |||||  |     | |||   || ||
Db        677 AAATCGGCAAGACACCTCAAAAGGATCTGCAGGCAGAAAAATCAACCTA 725


RESULT 14
```

```
US-09-790-988-1/c
; Sequence 1, Application US/09790988
; Patent No. US20020127687A1
; GENERAL INFORMATION:
;  APPLICANT: SHIGENOBU, SHUJI
;  APPLICANT:  WATANABE, HIDEMI
;  APPLICANT:  HATTORI, MASAHIRA
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
;  FILE REFERENCE: 081356/0159
;  CURRENT APPLICATION NUMBER: US/09/790,988
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: JP2000-107160
;  PRIOR FILING DATE: 2000-04-07
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 640681
;   TYPE: DNA
;   ORGANISM: Buchnera sp.
   09-790-988-1

  Query Match             13.3%;  Score 114.2;  DB 10;  Length 640681;
  Best Local Similarity   51.4%;  Pred. No. 4.2e-14;
  Matches  319;  Conservative    0;  Mismatches  293;  Indels    9;  Gaps    2;

Qy         88 CTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCGACCAGTTCTG 147
Db     511231 CTTTTAAAAGCAATGAAATATAGTGTATTTTCAGGTAGTAAAAGATTGCGTTCATCTTTA 511172

Qy        148 TTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAAGAGCGCAATT 207
Db     511171 ATATATTCAACTGGAGATGTTTTTAAAGTTAATATTACTACATTAGATGTTATATCTACT 511112

Qy        208 GCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAGCGATGGATAAT 267
Db     511111 GCCATCGAATTTATTCATTCATACTCTTTAATACATGATGACTTACCCTGTATGGATAAT 511052

Qy        268 GATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTGGACTGCGATA 327
Db     511051 GATAATTTTAGAAGAGGAAAGATATCTTGTCATGTAAAATACGGTGAAAGTACTTCTTTA 510992

Qy        328 TTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGATGAT------ 381
Db     510991 CTTGCTGGTGATGCTTTGCAAAGTCTTGCGTTTAATATTCTATCAAACAGTTTTATGCCA 510932

Qy        382 AGATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGCAAGTGGTCAT 441
Db     510931 AATGTATCTAATTTAAAACGTATAAAAATGATTTCTGAATTATCTTATTCCATTGGTTCA 510872

Qy        442 GTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACCAATTGATCTT 501
Db     510871 TCTGGGATGTGCATGGGTCAAAATTTAGATTTAGAAGCAGAAAAAAAAAGATGTAAATTTA 510812

Qy        502 GAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTTTTGCGGTTATG 561
Db     510811 TCTGAATTAGAAATAATTAATTTATATAAAACTTCTTTTTTTAATGCGTTCTGCTGTACGT 510752

Qy        562 AGTGCAGCAGATATCGCTAATGTCGATGATACAACTA---AAGAACATTTAGAAAGTTAT 618
Db     510751 TTAGTATATTTTTCTTCTAATAATTTTTCTAAATCTATATTATCAATATTAGATCTTTTT 510692

Qy        619 AGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTGCTATGGTGAT 678
Db     510691 TCGATTTCTATTGGTTTGGCATTTCAAATTCAGGATGATATTTTAGATTTTAAAAAAGAT 510632

Qy        679 GAAGCAAAGTTAGGTAAAAAA 699
Db     510631 AGTGTAAAAACAGACAATAAA 510611


RESULT 15
US-09-815-242-5972
; Sequence 5972, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5972
;   LENGTH: 900
;   TYPE: DNA
;   ORGANISM: Escherichia coli
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(900)
US-09-815-242-5972

  Query Match             11.9%;  Score 102.4;  DB 10;  Length 900;
  Best Local Similarity   47.9%;  Pred. No. 1.2e-12;
  Matches  363;  Conservative    0;  Mismatches  386;  Indels    9;  Gaps    2;

Qy         78 GGATACTCAGCTAGAAGAAAGTATGTTGTATTCATTAAATGCTGGAGGTAAACGCATCCG 137
Db         84 GAACACTCCCGTGGTCGAAACCATGCAGTATGGCGCATTATTAGGTGGTAAGCGCCTGCG 143

Qy        138 ACCAGTTCTGTTATTACTCACTTTAGATTCACTAAATACCGAGTATGAGTTAGGTATGAA 197
Db        144 ACCTTTCCTGGTTTATGCCACCGGTCATATGTTCGGCGTTAGCACAAACACGCTGGACGC 203

Qy        198 GAGCGCAATTGCACTAGAAATGATTCATACATATTCACTTATTCATGATGACCTACCAGC 257
Db        204 ACCCGCTGCCGCCGTTGAGTGTATCCACGCTTACTCATTAATTCATGATGATTTACCGGC 263

Qy        258 GATGGATAATGATGATTATCGACGAGGAAAATTAACAAATCATAAAGTATATGGTGAGTG 317
Db        264 AATGGATGATGACGATCTGCGTCGCGGTTTGCCAACCTGCCATGTGAAGTTTGGCGAAGC 323

Qy        318 GACTGCGATATTAGCAGGTGATGCTTTATTAACTAAAGCATTTGAACTTATTTCAAGTGA 377
Db        324 AAACGCGATTCTCGCTGGCGACGCTTTACAAACGCTGGCGTTCTCGATTTTAAGCGATGC 383

Qy        378 TGATA------GATTAACTGATGAAGTAAAAATAAAAGTTCTACAACGGCTGTCAATAGC 431
Db        384 CGATATGCCGGAAGTGTCGGACCGCGACAGAATTTCGATGATTTCTGAACTGGCGAGCGC 443

Qy        432 AAGTGGTCATGTTGGAATGGTCGGCGGTCAAATGTTAGATATGCAAAGCGAAGGCCAACC 491
Db        444 CAGTGGTATTGCCGGAATGTGCGGTGGTCAGGCATTAGATTTAGACGCGGAAGGCAAACA 503
```

```
Qy        492 AATTGATCTTGAAACTTTGGAAATGATACACAAAACAAAAACAGGAGCATTATTAACTTT 551
Db        504 CGTACCTCTGGACGCGCTTGAGCGTATTCATCGTCATAAAACCGGCGCATTGATTCGCGC 563

Qy        552 TGCGGTT---ATGAGTGCAGCAGATATCGCTAATGTCGATGATACAACTAAAGAACATTT 608
Db        564 CGCCGTTCGCCTTGGTGCATTAAGCGCCGGAGATAAAGGACGTCGTGCTCTGCCGGTACT 623

Qy        609 AGAAAGTTATAGTTATCATTTAGGTATGATGTTCCAGATTAAAGATGATTTATTAGACTG 668
Db        624 CGACAAGTATGCAGAGAGCATCGGCCTTGCCTTCCAGGTTCAGGATGACATCCTGGATGT 683

Qy        669 CTATGGTGATGAAGCAAAGTTAGGTAAAAAAGTGGGCAGCGATCTTGAAAATAATAAAAG 728
Db        684 GGTGGGAGATACTGCAACGTTGGGAAAACGCCAGGGTGCCGACCAGCAACTTGGTAAAAG 743

Qy        729 TACGTACGTGAGTTTATTAGGGAAAGATGGCGCAGAAGATAAATTGACTTATCATAGAGA 788
Db        744 TACCTACCCTGCACTTCTGGGTCTTGAGCAAGCCCGGAAGAAAGCCCGGGATCTGATCGA 803

Qy        789 CGCAGCAGTGGATGAACTAACGCAAATTGATGAACAAT 826
Db        804 CGATGCCCGTCAGTCGCTGAAACAACTGGCTGAACAGT 841
```

Search completed: May 30, 2003, 18:21:16
Job time : 187 secs